GenCore version 5.1.6

DM protein - protein search, using sw model

Run on: April 14, 2004, 15:40:43 ; Search time 22 Seconds
(without alignments)
1145.158 Million cell updates/sec

Title: US-09-632-722-2
Perfect score: 2634
Sequence: 1 MGRPLHLVLLSASLAGLLLL.....RGLPKAKSHAPEVITSSPLK 488

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2634	100.0	488	4	US-09-367-777-44	Sequence 44, Appl
2	2634	100.0	488	4	US-09-367-791A-27	Sequence 27, Appl
3	2557	97.1	488	1	US-08-487-037-1	Sequence 1, Appli
4	2441	92.7	448	5	PCT-US92-10068-1	Sequence 1, Appli
5	2439	92.6	448	1	US-08-295-411-3	Sequence 3, Appli
6	2439	92.6	448	2	US-08-955-471-3	Sequence 3, Appli
7	2439	92.6	448	5	PCT-US92-10242-3	Sequence 3, Appli
8	2249.5	85.4	437	1	US-08-487-037-2	Sequence 2, Appli
9	2241.5	85.1	437	1	US-08-487-037-3	Sequence 3, Appli
10	1851	70.3	492	1	US-08-469-486-2	Sequence 2, Appli
11	1851	70.3	492	2	US-08-469-658-2	Sequence 2, Appli
12	1847	70.1	487	1	US-08-469-486-53	Sequence 53, Appl
13	1847	70.1	487	2	US-08-469-658-53	Sequence 53, Appl
14	1631	61.9	306	1	US-08-330-978-1	Sequence 1, Appli
15	1631	61.9	306	1	US-08-474-042-1	Sequence 1, Appli
16	1631	61.9	306	1	US-08-484-558-1	Sequence 1, Appli
17	1631	61.9	306	1	US-08-774-592-1	Sequence 1, Appli
18	1354	51.4	254	1	US-08-330-978-3	Sequence 3, Appli
19	1354	51.4	254	1	US-08-474-042-3	Sequence 3, Appli
20	1354	51.4	254	1	US-08-484-558-3	Sequence 3, Appli
21	1354	51.4	254	1	US-08-774-592-3	Sequence 3, Appli
22	1305.5	49.6	247	3	US-08-944-483-49	Sequence 49, Appl
23	1289	48.9	241	1	US-08-330-978-4	Sequence 4, Appli
24	1289	48.9	241	1	US-08-474-042-4	Sequence 4, Appli
25	1289	48.9	241	1	US-08-484-558-4	Sequence 4, Appli
26	1289	48.9	241	1	US-08-774-592-4	Sequence 4, Appli
27	1051.5	39.9	461	6	5521070-2	Patent No. 5521070
28	1041.5	39.5	461	3	US-08-742-877-2	Sequence 2, Appli
29	1039.5	39.5	461	4	US-09-053-871A-21	Sequence 21, Appl
30	1039.5	39.5	461	4	US-10-133-907-5	Sequence 5, Appli
31	1002.5	38.1	415	1	US-08-073-531B-1	Sequence 1, Appli
32	1002.5	38.1	415	2	US-08-766-288-1	Sequence 1, Appli
33	996.5	37.8	415	4	US-09-118-748-2	Sequence 2, Appli
34	992.5	37.7	415	1	US-08-295-411-2	Sequence 2, Appli
35	992.5	37.7	415	2	US-08-955-471-2	Sequence 2, Appli
36	992.5	37.7	415	5	PCT-US92-10242-2	Sequence 2, Appli
37	952	36.1	444	1	US-08-475-845-2	Sequence 2, Appli
38	952	36.1	444	2	US-08-327-690-2	Sequence 2, Appli
39	952	36.1	444	2	US-08-660-289-2	Sequence 2, Appli
40	952	36.1	444	2	US-08-537-807-2	Sequence 2, Appli
41	952	36.1	444	2	US-08-871-003-2	Sequence 2, Appli
42	952	36.1	444	3	US-08-464-233-2	Sequence 2, Appli
43	952	36.1	444	3	US-09-189-607-2	Sequence 2, Appli
44	952	36.1	444	3	US-09-378-907-2	Sequence 2, Appli
45	952	36.1	444	5	PCT-US94-05779-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-367-777-44
; Sequence 44, Application US/09367777
; Patent No. 6562598
; GENERAL INFORMATION:
; APPLICANT: Himmelspach, Michele
; Pfleiderer, Michael
; Falkner, Falko-Guenter
; Eibl, Johann
; Dorner, Friedrich
; Schlokat, Uwe
; TITLE OF INVENTION: Factor X Deletion Mutants
; and Analogues Thereof
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/367,777
; FILING DATE: 10-No. 6562598-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AT A 336/97
; FILING DATE: 27-FEB-1997
; APPLICATION NUMBER: WO PCT/AT98/00046
; FILING DATE: 27-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ausenhus, Scott L.
; REGISTRATION NUMBER: 42,271
; REFERENCE/DOCKET NUMBER: 20695D-000900US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```
;        SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-367-777-44

  Query Match             100.0%;  Score 2634;  DB 4;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-203;
  Matches  488;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480

Qy        481 VITSSPLK 488
              ||||||||
Db        481 VITSSPLK 488


RESULT 2
US-09-367-791A-27
; Sequence 27, Application US/09367791A
; Patent No. 6573071
;   GENERAL INFORMATION:
;        APPLICANT: Himmelspach, Michele
;                   Schlokat, Uwe
;                   Dorner, Friedrich
;                   Fisch, Andreas
;                   Eibl, Johann
;        TITLE OF INVENTION: Factor X Analogues With
;                            a Modified Protease Cleavage Site
;        NUMBER OF SEQUENCES: 122
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/367,791A
;             FILING DATE: 12-No. 6573071-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: AT A 335/97
;             FILING DATE: 27-FEB-1997
;             APPLICATION NUMBER: WO PCT/AT98/00045
;             FILING DATE: 27-FEB-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ausenhus, Scott L.
;             REGISTRATION NUMBER: 42,471
;             REFERENCE/DOCKET NUMBER: 20695D-000700US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 27:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 488 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-367-791A-27

  Query Match             100.0%;  Score 2634;  DB 4;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-203;
  Matches  488;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480

Qy        481 VITSSPLK 488
              ||||||||
Db        481 VITSSPLK 488


RESULT 3
US-08-487-037-1
; Sequence 1, Application US/08487037
; Patent No. 5795863
;  GENERAL INFORMATION:
;    APPLICANT:  Wolf, David L.
;    TITLE OF INVENTION:  RECOMBINANT AGENTS AFFECTING THROMBOSIS
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,037
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Reid G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0002.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  488 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  both
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  -17
;      OTHER INFORMATION:  /note= "Location of Intron A"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  (37^38)
;      OTHER INFORMATION:  /note= "Location of Intron B"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  46
;      OTHER INFORMATION:  /note= "Location of Intron C"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  63
;      OTHER INFORMATION:  /note= "Amino acid represented by
;      OTHER INFORMATION:  the greek letter Beta"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  84
;      OTHER INFORMATION:  /note= "Location of Intron D"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  128
;      OTHER INFORMATION:  /note= "Location of Intron E"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  (209^210)
;      OTHER INFORMATION:  /note= "Location of Intron F"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  249
;      OTHER INFORMATION:  /note= "Location of Intron G"
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  -40..0
;      OTHER INFORMATION:  /note= "Pre-Pro leader sequence"
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..139
;      OTHER INFORMATION:  /note= "Factor Xa- Light chain"
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  143..194
;      OTHER INFORMATION:  /note= "Activation Peptide"
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  195..448
;      OTHER INFORMATION:  /note= "Factor Xa-Heavy Chain"
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  group(17..22, 50..61, 55..70, 72..81, 89..100, 96
;      LOCATION:  ..109, 111..124, 132..302, 201..206, 221..237,
;      LOCATION:  350..364, 375..403)
US-08-487-037-1

  Query Match             97.1%;  Score 2557;  DB 1;  Length 488;
  Best Local Similarity   97.3%;  Pred. No. 5.3e-197;
  Matches  475;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||  |||||| | ||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLTTMKKGHLTRTCMTT 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||  || || |||||| ||||||||||||||||||||| ||||||||||||||||||
Db         61 TCSYTTARTVFTDSDKTNTFWNKYKDGDQCETSPCQNQGKCKXGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        421 GPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480

Qy        481 VITSSPLK 488
              ||||||||
Db        481 VITSSPLK 488


RESULT 4
PCT-US92-10068-1
; Sequence 1, Application PC/TUS9210068
;  GENERAL INFORMATION:
;    APPLICANT:  Altieri, Dario C
;    APPLICANT:  Edgington, Thomas S
;    APPLICANT:  Fair, Daryl S
;    TITLE OF INVENTION:  Factor X-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Inflammation
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 North Torrey Pines Road
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/10068
;       FILING DATE:  19921120
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/798,221
;       FILING DATE:  22-NOV-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fitting, Thomas
;       REGISTRATION NUMBER:  SCR1221P
;       REFERENCE/DOCKET NUMBER:  34,163
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-554-2937
;       TELEFAX:  619-554-6312
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  448 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  1..139
;       OTHER INFORMATION:  /note= "Factor X Light Chain"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  140..142
;       OTHER INFORMATION:  /note= "Factor X Connecting
;       OTHER INFORMATION:  Tripeptide"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  143..448
;       OTHER INFORMATION:  /note= "Factor X Heavy Chain"
PCT-US92-10068-1

  Query Match             92.7%;  Score 2441;  DB 5;  Length 448;
  Best Local Similarity   99.8%;  Pred. No. 9.4e-188;
  Matches  447;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 60

Qy        101 CKDGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CKDGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 120

Qy        161 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 180

Qy        221 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 240

Qy        281 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 300

Qy        341 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 360

Qy        401 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLK 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLK 420

Qy        461 WIDRSMKTRGLPKAKSHAPEVITSSPLK 488
              ||||||||||||||||||||||||||||
Db        421 WIDRSMKTRGLPKAKSHAPEVITSSPLK 448


RESULT 5
US-08-295-411-3
; Sequence 3, Application US/08295411
; Patent No. 5679639
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5679639th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,411
;      FILING DATE:  22-AUG-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/793,989
;      FILING DATE:  18-NOV-1991
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  448 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..139
;      OTHER INFORMATION:  /note= "Factor X Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  140..142
;      OTHER INFORMATION:  /note= "Factor X Connecting
;      OTHER INFORMATION:  Tripeptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  143..448
;      OTHER INFORMATION:  /note= "Factor X Heavy Chain"
US-08-295-411-3

  Query Match             92.6%;  Score 2439;  DB 1;  Length 448;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-187;
  Matches  446;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 60

Qy        101 CKDGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 160
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CKBGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 120

Qy        161 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 180

Qy        221 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 240

Qy        281 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 300

Qy        341 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 360

Qy        401 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLK 460
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        361 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLK 420

Qy        461 WIDRSMKTRGLPKAKSHAPEVITSSPLK 488
              ||||||||||||||||||||||||||||
Db        421 WIDRSMKTRGLPKAKSHAPEVITSSPLK 448


RESULT 6
US-08-955-471-3
; Sequence 3, Application US/08955471
; Patent No. 5968751
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf M.
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 No. 5968751th Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/955,471
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/295,411
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  TSRI263.0C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  448 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..139
;      OTHER INFORMATION:  /note= "Factor X Light Chain"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  140..142
;      OTHER INFORMATION:  /note= "Factor X Connecting
;      OTHER INFORMATION:  Tripeptide"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  143..448
;      OTHER INFORMATION:  /note= "Factor X Heavy Chain"
US-08-955-471-3

  Query Match             92.6%;  Score 2439;  DB 2;  Length 448;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-187;
  Matches  446;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 60

Qy        101 CKDGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 160
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CKBGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 120

Qy        161 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 180

Qy        221 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 240

Qy        281 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 300

Qy        341 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 360

Qy        401 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLK 460
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        361 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLK 420

Qy        461 WIDRSMKTRGLPKAKSHAPEVITSSPLK 488
              ||||||||||||||||||||||||||||
Db        421 WIDRSMKTRGLPKAKSHAPEVITSSPLK 448


RESULT 7
PCT-US92-10242-3
; Sequence 3, Application PC/TUS9210242
;  GENERAL INFORMATION:
;    APPLICANT:  Griffin, John H.
;    APPLICANT:  Mesters, Rolf
;    TITLE OF INVENTION:  Serine Protease-Derived Polypeptides and
;    TITLE OF INVENTION:  Anti-Peptide Antibodies, Systems and Therapeutic Methods
;    TITLE OF INVENTION:  for Inhibiting Coagulation
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Office of Patent Counsel, The Scripps
;      ADDRESSEE:  Research Institute
;      STREET:  10666 North Torrey Pines Road, TPC 8
;      CITY:  La Jolla
;      STATE:  CA
```

```
;       COUNTRY:  USA
;       ZIP:  92037
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/10242
;       FILING DATE:  19921118
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/793,989
;       FILING DATE:  18-NOV-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fitting, Thomas
;       REGISTRATION NUMBER:  34,163
;       REFERENCE/DOCKET NUMBER:  SCR0472P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-554-2937
;       TELEFAX:  619-554-6312
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  448 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  1..139
;       OTHER INFORMATION:  /note= "Factor X Light Chain"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  140..142
;       OTHER INFORMATION:  /note= "Factor X Connecting
;       OTHER INFORMATION:  Tripeptide"
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  143..448
;       OTHER INFORMATION:  /note= "Factor X Heavy Chain"
PCT-US92-10242-3

  Query Match             92.6%;  Score 2439;  DB 5;  Length 448;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-187;
  Matches  446;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 60

Qy        101 CKDGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 160
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CKBGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 120

Qy        161 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 180

Qy        221 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 240

Qy        281 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAP 300

Qy        341 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQ 360

Qy        401 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLK 460
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        361 NMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLK 420

Qy        461 WIDRSMKTRGLPKAKSHAPEVITSSPLK 488
              ||||||||||||||||||||||||||||
Db        421 WIDRSMKTRGLPKAKSHAPEVITSSPLK 448


RESULT 8
US-08-487-037-2
; Sequence 2, Application US/08487037
; Patent No. 5795863
;  GENERAL INFORMATION:
;    APPLICANT:  Wolf, David L.
;    TITLE OF INVENTION:  RECOMBINANT AGENTS AFFECTING THROMBOSIS
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,037
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Reid G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0002.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  437 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  both
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..139
;      OTHER INFORMATION:  /note= "Factor Xa-Light Chain"
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  -40..0
;      OTHER INFORMATION:  /note= "Pre-Pro leader sequence"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  -17
;      OTHER INFORMATION:  /note= "Location of Intron A"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  (37^38)
;      OTHER INFORMATION:  /note= "Location of Intron B"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  46
;      OTHER INFORMATION:  /note= "Location of Intron C"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  63
;      OTHER INFORMATION:  /note= "An amino acid represented
;      OTHER INFORMATION:  by the greek letter Beta"
;    FEATURE:
;      NAME/KEY:  Modified-site
```

```
;       LOCATION:  84
;       OTHER INFORMATION:  /note= "Location of Intron D"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  128
;       OTHER INFORMATION:  /note= "Location of Intron E"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  (158^159)
;       OTHER INFORMATION:  /note= "Location of Intron F"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  198
;       OTHER INFORMATION:  /note= "Location of Intron G"
;     FEATURE:
;       NAME/KEY:  Disulfide-bond
;       LOCATION:  group(17..22, 50..61, 55..70, 72..81, 89..100, 96
;       LOCATION:  ..109, 111..124, 132..251, 150..155, 170..186,
;       LOCATION:  299..313, 324..352)
;S-08-487-037-2

  Query Match             85.4%;  Score 2249.5;  DB 1;  Length 437;
  Best Local Similarity   86.9%;  Pred. No. 2e-172;
  Matches  424;  Conservative    0;  Mismatches   13;  Indels   51;  Gaps    1;

;y          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
;b          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLTTMKKGHLTRTCMTT 60

;y         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||  ||  ||  ||||||  ||||||||||||||||||||||||  ||||||||||||||||
;b         61 TCSYTTARTVFTDSDKTNTFWNKYKDGDQCETSPCQNQGKCKXGLGEYTCTCLEGFEGKN 120

;y        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
;b        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

;y        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||                                                  |||||||
;b        181 KR----------------------------------------------------RIVGGQE 189

;y        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
;b        190 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 249

;y        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
;b        250 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 309

;y        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
;b        310 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 369

;y        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
;b        370 GPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 429

;y        481 VITSSPLK 488
              ||||||||
;b        430 VITSSPLK 437


ESULT 9
S-08-487-037-3
 Sequence 3, Application US/08487037
 Patent No. 5795863
  GENERAL INFORMATION:
    APPLICANT:  Wolf, David L.
    TITLE OF INVENTION:  RECOMBINANT AGENTS AFFECTING THROMBOSIS
    NUMBER OF SEQUENCES:  11
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  MORRISON & FOERSTER
      STREET:  2000 Pennsylvania Avenue, NW
;       CITY:  Washington
;       STATE:  DC
;       COUNTRY:  USA
;       ZIP:  20006-1812
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/487,037
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Adler, Reid G.
;       REGISTRATION NUMBER:  30,988
;       REFERENCE/DOCKET NUMBER:  2803-0002.02
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 887-1500
;       TELEFAX:  (202) 887-0763
;       TELEX:  90-4030
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  437 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  both
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  -40..397
;       OTHER INFORMATION:  /note= "Same features apply from
;       OTHER INFORMATION:  SEQ ID NO:2"
;     FEATURE:
;       NAME/KEY:  Protein
;       LOCATION:  1..139
;       OTHER INFORMATION:  /note= "Factor Xa - Light Chain"
;     FEATURE:
;       NAME/KEY:  Peptide
;       LOCATION:  -40..0
;       OTHER INFORMATION:  /note= "Pre-Pro leader sequence"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  -17
;       OTHER INFORMATION:  /note= "Location of Intron A"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  (37^38)
;       OTHER INFORMATION:  /note= "Location of Intron B"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  46
;       OTHER INFORMATION:  /note= "Location of Intron C"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  63
;       OTHER INFORMATION:  /note= "An amino acid represented
;       OTHER INFORMATION:  by the greek letter Beta"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  84
;       OTHER INFORMATION:  /note= "Location of Intron D"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  (158^159)
;       OTHER INFORMATION:  /note= "Location of Intron F"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  198
;       OTHER INFORMATION:  /note= "Location of Intron G"
;     FEATURE:
;       NAME/KEY:  Disulfide-bond
;       LOCATION:  group(17..22, 50..61, 55..70, 72..81, 89..100, 96
;       LOCATION:  ..109, 111..124, 132..251, 150..155, 170..186,
```

```
;      LOCATION:  299..313, 324..352)
US-08-487-037-3

  Query Match             85.1%;  Score 2241.5;  DB 1;  Length 437;
  Best Local Similarity   86.5%;  Pred. No. 8.9e-172;
  Matches  422;  Conservative    2;  Mismatches   13;  Indels   51;  Gaps    1;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||  |||||| | ||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLTTMKKGHLTRTCMTT 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||  || || |||||| ||||||||||||||||||||||| ||||||||||||||||||
Db         61 TCSYTTARTVFTDSDKTNTFWNKYKDGDQCETSPCQNQGKCKXGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||                                                  |||||||
Db        181 KR----------------------------------------------------RIVGGQE 189

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 249

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        250 AVHEVEVVIKHNRFTKETYDFNIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 309

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        310 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDAG 369

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        370 GPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 429

Qy        481 VITSSPLK 488
              ||||||||
Db        430 VITSSPLK 437


RESULT 10
US-08-469-486-2
; Sequence 2, Application US/08469486
; Patent No. 5739281
;  GENERAL INFORMATION:
;    APPLICANT:  Thoegersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  Improved method for the refolding of
;    TITLE OF INVENTION:  proteins
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,486
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  492 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-469-486-2

  Query Match             70.3%;  Score 1851;  DB 1;  Length 492;
  Best Local Similarity   69.8%;  Pred. No. 2.1e-140;
  Matches  340;  Conservative   58;  Mismatches   81;  Indels    8;  Gaps    4;

Qy          5 LHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEETCSY 64
              |||||||  :| |||     |:|: |:||: :| |  ||||||||:|:|:|||||:|| ||
Db          5 LHLVLLSTALGGLLRPAGSVFLPRDQAHRVLQRARRANSFLEEVKQGNLERECLEEACSL 64

Qy         65 EEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELF 124
              |||||||||:::|:|||:|||||||||  || ||| ||||:|:||||| |||||||||
Db         65 EEAREVFEDAEQTDEFWSKYKDGDQCEGHPCLNQGHCKDGIGDYTCTCAEGFEGKNCEFS 124

Qy        125 TRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSV 184
              ||::|||||| ||||| ||:: | |||| || | |: |:|: |  :|||| |  | :|
Db        125 TREICSLDNGGCDQFCREERSEVRCSCAHGYVLGDDSKSCVSTERFPCGKFTQGRSRR-- 182

Qy        185 AQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGD--NNLTRIVGGQECK 242
                |  :| :| |:   : || ||| |||:  |||  |:|:|  |:  : : |||||::|
Db        183 -WAIHTSEDALDASELEHYDPADLSPTESSLDLLGLNRTEPSAGEDGSQVVRIVGGRDCA 241

Qy        243 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 302
              :|||||||||:|||||||||||||:|||:|||||||:||||| |||||||||||| |
Db        242 EGECPWQALLVNEENEGFCGGTILNEFYVLTAAHCLHQAKRFTVRVGDRNTEQEEGNEMA 301

Qy        303 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 362
              ||||: :||:|| |||||||||||||| || |||||||||:||||:||||||||||
Db        302 HEVEMTVKHSRFVKETYDFDIAVLRLKTPIRFRRNVAPACLPEKDWAEATLMTQKTGIVS 361

Qy        363 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 422
              |||||||||| |: |||||||||::|||||| || |||||||||: ||||||||||
Db        362 GFGRTHEKGRLSSTLKMLEVPYVDRSTCKLSSSFTITPNMFCAGYDTQPEDACQGDSGGP 421

Qy        423 HVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTR-GLPKAKSH--AP 479
              |||||||||||||||||| ||||||:|:||||: ||||||: || | |   :: |  ||
Db        422 HVTRFKDTYFVTGIVSWGEGCARKGKFGVYTKVSNFLKWIDKIMKARAGAAGSRGHSEAP 481

Qy        480 EVITSSP 486
                | |
Db        482 ATWTVPP 488


RESULT 11
US-08-469-658-2
; Sequence 2, Application US/08469658
; Patent No. 5917018
;  GENERAL INFORMATION:
;    APPLICANT:  Th egersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  IMPROVED METHOD FOR THE REFOLDING OF
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
```

```
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,658
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  492 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-469-658-2

  Query Match             70.3%;  Score 1851;  DB 2;  Length 492;
  Best Local Similarity   69.8%;  Pred. No. 2.1e-140;
  Matches  340;  Conservative   58;  Mismatches   81;  Indels    8;  Gaps    4;

Qy          5 LHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEETCSY 64
              ||||||| :| |||    |:|: |:||: :| |  ||||||||:|:|:|||||:|| ||
Db          5 LHLVLLSTALGGLLRPAGSVFLPRDQAHRVLQRARRANSFLEEVKQGNLERECLEEACSL 64

Qy         65 EEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELF 124
              |||||||||:::|:|||:|||||||||  || ||| ||||:|:||||| |||||||||
Db         65 EEAREVFEDAEQTDEFWSKYKDGDQCEGHPCLNQGHCKDGIGDYTCTCAEGFEGKNCEFS 124

Qy        125 TRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSV 184
              ||::|||||| ||||| ||:: | |||| || | |: |:|: |  :|||| |  | :|
Db        125 TREICSLDNGGCDQFCREERSEVRCSCAHGYVLGDDSKSCVSTERFPCGKFTQGRSRR-- 182

Qy        185 AQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGD--NNLTRIVGGQECK 242
                |  :| :| |:   : || ||| |||:  |||  |:|:|  |:  : : ||||::|
Db        183 -WAIHTSEDALDASELEHYDPADLSPTESSLDLLGLNRTEPSAGEDGSQVVRIVGGRDCA 241

Qy        243 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 302
              :|||||||||:|||||||||||||:|||:|||||||:||||| |||||||||||||| |
Db        242 EGECPWQALLVNEENEGFCGGTILNEFYVLTAAHCLHQAKRFTVRVGDRNTEQEEGNEMA 301

Qy        303 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 362
              ||||: :||:|| |||||||||||||||| || |||||||||:||||:|||||||||||
Db        302 HEVEMTVKHSRFVKETYDFDIAVLRLKTPIRFRRNVAPACLPEKDWAEATLMTQKTGIVS 361

Qy        363 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 422
              |||||||||| |: |||||||||||::|||||| || |||||||||: ||||||||||
Db        362 GFGRTHEKGRLSSTLKMLEVPYVDRSTCKLSSSFTITPNMFCAGYDTQPEDACQGDSGGP 421

Qy        423 HVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTR-GLPKAKSH--AP 479
              |||||||||||||||| |||||:|:||||: |||||: || | |   :: |  ||
Db        422 HVTRFKDTYFVTGIVSWGEGCARKGKFGVYTKVSNFLKWIDKIMKARAGAAGSRGHSEAP 481

Qy        480 EVITSSP 486
```

```
                | |
Db        482 ATWTVPP 488


RESULT 12
US-08-469-486-53
; Sequence 53, Application US/08469486
; Patent No. 5739281
;  GENERAL INFORMATION:
;    APPLICANT:  Thoegersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  Improved method for the refolding of
;    TITLE OF INVENTION:  proteins
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,486
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  487 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-469-486-53

  Query Match             70.1%;  Score 1847;  DB 1;  Length 487;
  Best Local Similarity   70.4%;  Pred. No. 4.3e-140;
  Matches  336;  Conservative   59;  Mismatches   76;  Indels    6;  Gaps    3;

Qy          5 LHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEETCSY 64
              ||||||| :| |||    |:|: |:||: :| |  ||||||||:|:|:|||||:|| ||
Db          5 LHLVLLSTALGGLLRPAGSVFLPRDQAHRVLQRARRANSFLEEVKQGNLERECLEEACSL 64

Qy         65 EEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELF 124
              |||||||||:::|:|||:|||||||||  || ||| ||||:|:||||| |||||||||
Db         65 EEAREVFEDAEQTDEFWSKYKDGDQCEGHPCLNQGHCKDGIGDYTCTCAEGFEGKNCEFS 124

Qy        125 TRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSV 184
              ||::|||||| ||||| ||:: | |||| || | |: |:|: |  :|||| |  | :|
Db        125 TREICSLDNGGCDQFCREERSEVRCSCAHGYVLGDDSKSCVSTERFPCGKFTQGRSRR-- 182

Qy        185 AQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGD--NNLTRIVGGQECK 242
                |  :| :| |:   : || ||| |||:  |||  |:|:|  |:  : : ||||::|
Db        183 -WAIHTSEDALDASELEHYDPADLSPTESSLDLLGLNRTEPSAGEDGSQVVRIVGGRDCA 241

Qy        243 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 302
```

```
Db        242 EGECPWQALLVNEENEGFCGGTILNEFYVLTAAHCLHQAKRFTVRVGDRNTEQEEGNEMA 301

Qy        303 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 362

Db        302 HEVEMTVKHSRFVKETYDFDIAVLRLKTPIRFRRNVAPACLPEKDWAEATLMTQKTGIVS 361

Qy        363 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 422

Db        362 GFGRTHEKGRLSSTLKMLEVPYVDRSTCKLSSSFTITPNMFCAGYDTQPEDACQGDSGGP 421

Qy        423 HVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTR-GLPKAKSHA 478

Db        422 HVTRFKDTYPVTGIVSWGEGCARKGKFGVYTKVSNFLKWIDKIMKARAGAAGSRGHS 478


RESULT 13
US-08-469-658-53
; Sequence 53, Application US/08469658
; Patent No. 5917018
;  GENERAL INFORMATION:
;    APPLICANT:  Th egersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  IMPROVED METHOD FOR THE REFOLDING OF
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,658
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  487 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-469-658-53

  Query Match             70.1%;  Score 1847;  DB 2;  Length 487;
  Best Local Similarity   70.4%;  Pred. No. 4.3e-140;
  Matches  336;  Conservative   59;  Mismatches   76;  Indels    6;  Gaps    3;

Qy          5 LHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEETCSY 64

Db          5 LHLVLLSTALGGLLRPAGSVFLPRDQAHRVLQRARRANSFLEEVKQGNLERECLEEACSL 64

Qy         65 EEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELF 124

Db         65 EEAREVFEDAEQTDEFWSKYKDGDQCEGHPCLNQGHCKDGIGDYTCTCAEGFEGKNCEFS 124

Qy        125 TRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSV 184

Db        125 TREICSLDNGGCDQFCREERSEVRCSCAHGYVLGDDSKSCVSTERFPCGKFTQGRSRR-- 182

Qy        185 AQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGD--NNLTRIVGGQECK 242

Db        183 -WAIHTSEDALDASELEHYDPADLSPTESSLDLLGLNRTEPSAGEDGSQVVRIVGGRDCA 241

Qy        243 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 302

Db        242 EGECPWQALLVNEENEGFCGGTILNEFYVLTAAHCLHQAKRFTVRVGDRNTEQEEGNEMA 301

Qy        303 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 362

Db        302 HEVEMTVKHSRFVKETYDFDIAVLRLKTPIRFRRNVAPACLPEKDWAEATLMTQKTGIVS 361

Qy        363 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 422

Db        362 GFGRTHEKGRLSSTLKMLEVPYVDRSTCKLSSSFTITPNMFCAGYDTQPEDACQGDSGGP 421

Qy        423 HVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTR-GLPKAKSHA 478

Db        422 HVTRFKDTYPVTGIVSWGEGCARKGKFGVYTKVSNFLKWIDKIMKARAGAAGSRGHS 478


RESULT 14
US-08-330-978-1
; Sequence 1, Application US/08330978
; Patent No. 5589571
;  GENERAL INFORMATION:
;    APPLICANT:  King, Robert
;    TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF INHIBITED
;    TITLE OF INVENTION:  FORMS OF ACTIVATED BLOOD FACTORS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/330,978
;      FILING DATE:  28-OCT-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/484,558
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  2803-0007.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)887-1500
;      TELEFAX:  (202)822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  306 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
```

```
;      LOCATION:  59..64
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  79..95
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  160
;      OTHER INFORMATION:  /note= "Disulfide linkage to
;      OTHER INFORMATION:  residue 132 of SEQ ID NO:2"
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  208..222
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  233..261
US-08-330-978-1

  Query Match             61.9%;  Score 1631;  DB 1;  Length 306;
  Best Local Similarity   99.7%;  Pred. No. 4.9e-123;
  Matches  305;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        183 SVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECK 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECK 60

Qy        243 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 120

Qy        303 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 180

Qy        363 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 240

Qy        423 HVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVI 482
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 HVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVI 300

Qy        483 TSSPLK 488
              ||||||
Db        301 TSSPLK 306


RESULT 15
US-08-474-042-1
; Sequence 1, Application US/08474042
; Patent No. 5589572
;  GENERAL INFORMATION:
;    APPLICANT:  King, Robert
;    TITLE OF INVENTION:  PROCESS FOR PRODUCTION OF INHIBITED
;    TITLE OF INVENTION:  FORMS OF ACTIVATED BLOOD FACTORS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,042
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/484,558
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  2803-0007.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)887-1500
;      TELEFAX:  (202)822-0168
;      TELEX:  90-4030 MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  306 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  59..64
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  79..95
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  160
;      OTHER INFORMATION:  /note= "Disulfide linkage to
;      OTHER INFORMATION:  residue 132 of SEQ ID NO:2"
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  208..222
;    FEATURE:
;      NAME/KEY:  Disulfide-bond
;      LOCATION:  233..261
US-08-474-042-1

  Query Match             61.9%;  Score 1631;  DB 1;  Length 306;
  Best Local Similarity   99.7%;  Pred. No. 4.9e-123;
  Matches  305;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        183 SVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECK 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECK 60

Qy        243 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAV 120

Qy        303 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVS 180

Qy        363 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGP 240

Qy        423 HVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVI 482
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        241 HVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVI 300

Qy        483 TSSPLK 488
              ||||||
Db        301 TSSPLK 306


Search completed: April 14, 2004, 15:41:22
Job time : 24 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 14, 2004, 15:40:44 ; Search time 47 Seconds
(without alignments)
2753.708 Million cell updates/sec

Title: US-09-632-722-2
Perfect score: 2634
Sequence: 1 MGRPLHLVLLSASLAGLLLL.....RGLPKAKSHAPEVITSSPLK 488

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1082010 seqs, 265213723 residues

Total number of hits satisfying chosen parameters: 1082010

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2634	100.0	488	14	US-10-348-504-44	Sequence 44, Appl
2	2634	100.0	488	14	US-10-407-123-27	Sequence 27, Appl
3	2630	99.8	488	12	US-10-406-031-27	Sequence 27, Appl
4	1646	62.5	309	15	US-10-360-101-233	Sequence 233, App
5	1251	47.5	467	12	US-10-406-031-5	Sequence 5, Appli
6	1238	47.0	467	12	US-10-406-031-2	Sequence 2, Appli
7	1238	47.0	467	12	US-10-406-031-8	Sequence 8, Appli
8	1223	46.4	455	12	US-10-406-031-17	Sequence 17, Appl
9	1222	46.4	454	12	US-10-406-031-11	Sequence 11, Appl
10	1216	46.2	453	12	US-10-406-031-14	Sequence 14, Appl
11	1135.5	43.1	376	12	US-10-406-031-31	Sequence 31, Appl
12	1041.5	39.5	461	9	US-09-884-901-3	Sequence 3, Appli
13	1041.5	39.5	461	14	US-10-234-406-8	Sequence 8, Appli
14	1041.5	39.5	461	16	US-10-038-854-92	Sequence 92, Appl
15	1041.5	39.5	461	16	US-10-239-498A-5	Sequence 5, Appli
16	1039.5	39.5	456	16	US-10-038-854-95	Sequence 95, Appl
17	1039.5	39.5	456	16	US-10-038-854-96	Sequence 96, Appl
18	1039.5	39.5	461	14	US-10-132-829-5	Sequence 5, Appli
19	1039.5	39.5	461	14	US-10-234-406-6	Sequence 6, Appli
20	1039.5	39.5	461	14	US-10-133-907-5	Sequence 5, Appli
21	1039.5	39.5	461	16	US-10-038-854-93	Sequence 93, Appl
22	1038.5	39.4	461	16	US-10-038-854-94	Sequence 94, Appl
23	1032	39.2	462	12	US-10-411-037-10	Sequence 10, Appl
24	1032	39.2	462	12	US-10-411-026-10	Sequence 10, Appl
25	1026	39.0	456	12	US-10-406-031-28	GENERAL INFORMATI
26	996.5	37.8	415	9	US-09-118-748-2	Sequence 2, Appli
27	974	37.0	421	12	US-10-406-031-30	Sequence 30, Appl
28	953	36.2	444	12	US-10-411-037-8	Sequence 8, Appli
29	953	36.2	444	12	US-10-382-248-34	Sequence 34, Appl
30	953	36.2	444	12	US-10-411-026-8	Sequence 8, Appli
31	948	36.0	466	14	US-10-017-122-2	Sequence 2, Appli
32	948	36.0	466	15	US-10-375-741-14	Sequence 14, Appl
33	909.5	34.5	406	10	US-09-782-587B-3	Sequence 3, Appli
34	909.5	34.5	406	15	US-10-383-898-1	Sequence 1, Appli
35	905.5	34.4	405	15	US-10-360-101-225	Sequence 225, App
36	865	32.8	394	16	US-10-038-854-6	Sequence 6, Appli
37	852.5	32.4	406	10	US-09-782-587B-1	Sequence 1, Appli
38	852.5	32.4	406	14	US-10-109-498-1	Sequence 1, Appli
39	852.5	32.4	406	14	US-10-255-032-1	Sequence 1, Appli
40	852.5	32.4	406	14	US-10-281-727-1	Sequence 1, Appli
41	852.5	32.4	406	15	US-10-386-898-7	Sequence 7, Appli
42	843.5	32.0	419	12	US-10-382-248-36	Sequence 36, Appl
43	827	31.4	461	10	US-09-978-917A-2	Sequence 2, Appli
44	827	31.4	461	12	US-09-997-623-2	Sequence 2, Appli
45	827	31.4	461	14	US-10-182-263-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-348-504-44
; Sequence 44, Application US/10348504
; Publication No. US20030138914A1
; GENERAL INFORMATION:
; APPLICANT: Himmelspach, Michele
; Pfleiderer, Michael
; Falkner, Falko-Guenter
; Eibl, Johann
; Dorner, Friedrich
; Schlokat, Uwe
; TITLE OF INVENTION: Factor X Deletion Mutants
; and Analogues Thereof
; NUMBER OF SEQUENCES: 145
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/348,504
; FILING DATE: 29-Jan-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/367,777
; FILING DATE: 10-No. US20030138914A1-1999
; APPLICATION NUMBER: AT A 336/97
; FILING DATE: 27-FEB-1997
; APPLICATION NUMBER: WO PCT/AT98/00046
; FILING DATE: 27-FEB-1998
; ATTORNEY/AGENT INFORMATION:

```
;             NAME: Ausenhus, Scott L.
;             REGISTRATION NUMBER: 42,271
;             REFERENCE/DOCKET NUMBER: 20695D-000900US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-576-0200
;             TELEFAX: 415-576-0300
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 44:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 488 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-10-348-504-44

  Query Match             100.0%;  Score 2634;  DB 14;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 8e-211;
  Matches  488;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480

Qy        481 VITSSPLK 488
              ||||||||
Db        481 VITSSPLK 488


RESULT 2
US-10-407-123-27
; Sequence 27, Application US/10407123
; Publication No. US20030181381A1
;   GENERAL INFORMATION:
;        APPLICANT: Himmelspach, Michele
;                   Schlokat, Uwe
;                   Dorner, Friedrich
;                   Fisch, Andreas
;                   Eibl, Johann
;        TITLE OF INVENTION: Factor X Analogues With
;                            a Modified Protease Cleavage Site
;        NUMBER OF SEQUENCES: 122
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/407,123
;             FILING DATE: 04-Apr-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/367,791A
;             FILING DATE: 12-No. US20030181381A1-1999
;             APPLICATION NUMBER: AT A 335/97
;             FILING DATE: 27-FEB-1997
;             APPLICATION NUMBER: WO PCT/AT98/00045
;             FILING DATE: 27-FEB-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ausenhus, Scott L.
;             REGISTRATION NUMBER: 42,471
;             REFERENCE/DOCKET NUMBER: 20695D-000700US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 27:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 488 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-10-407-123-27

  Query Match             100.0%;  Score 2634;  DB 14;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 8e-211;
  Matches  488;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480

Qy        481 VITSSPLK 488
              ||||||||
Db        481 VITSSPLK 488


RESULT 3
US-10-406-031-27
; Sequence 27, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
;  APPLICANT: Masci, Paul Pantaleone
;  APPLICANT:  De Jersey, John
;  APPLICANT:  Lavin, Martin
;  TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
;  FILE REFERENCE: 15685-002001
;  CURRENT APPLICATION NUMBER: US/10/406,031
;  CURRENT FILING DATE:  2003-04-02
;  PRIOR APPLICATION NUMBER: AU 2003901033
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: AU PS1483
;  PRIOR FILING DATE: 2002-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 488
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-406-031-27

  Query Match           99.8%;  Score 2630;  DB 12;  Length 488;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-210;
  Matches  487;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        421 GPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480

Qy        481 VITSSPLK 488
              ||||||||
Db        481 VITSSPLK 488


RESULT 4
US-10-360-101-233
; Sequence 233, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
;  APPLICANT: Moll, Gert  N.
;  APPLICANT:  Leenhouts, Cornelis J.
;  TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
;  FILE REFERENCE: 2183-5673
;  CURRENT APPLICATION NUMBER: US/10/360,101
;  CURRENT FILING DATE:  2003-02-07
;  PRIOR APPLICATION NUMBER: EP 02077060.8
;  PRIOR FILING DATE: 2002-05-24
;  NUMBER OF SEQ ID NOS: 309
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 233
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: heavy chain sequence of factor X
US-10-360-101-233

  Query Match           62.5%;  Score 1646;  DB 15;  Length 309;
  Best Local Similarity   99.7%;  Pred. No. 8.7e-129;
  Matches  308;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        180 RKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQ 239
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQ 60

Qy        240 ECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGG 299
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGG 120

Qy        300 EAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTG 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTG 180

Qy        360 IVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDS 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDS 240

Qy        420 GGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAP 479
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        241 GGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAP 300

Qy        480 EVITSSPLK 488
              |||||||||
Db        301 EVITSSPLK 309


RESULT 5
US-10-406-031-5
; Sequence 5, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
;  APPLICANT: Masci, Paul Pantaleone
;  APPLICANT:  De Jersey, John
;  APPLICANT:  Lavin, Martin
;  TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
;  FILE REFERENCE: 15685-002001
;  CURRENT APPLICATION NUMBER: US/10/406,031
;  CURRENT FILING DATE:  2003-04-02
;  PRIOR APPLICATION NUMBER: AU 2003901033
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: AU PS1483
;  PRIOR FILING DATE: 2002-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 467
;   TYPE: PRT
;   ORGANISM: Oxyuranus scutellatus
```

US-10-406-031-5

Query Match 47.5%; Score 1251; DB 12; Length 467;
Best Local Similarity 49.4%; Pred. No. 1.2e-95;
Matches 241; Conservative 68; Mismatches 141; Indels 38; Gaps 7;

```
Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              |    | | |:   |  |      ::|:: : ||  | |  ||||  || : |::||||:||
Db          1 MAPQLLLCLILTFLWSLPEAESNVFLKSKVANRFLQRTKRANSLYEEFRSGNIERECIEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
               || ||||||||| :||  ||| | ||||| ::||  :| ||||:| |||||| |:||||
Db         61 RCSKEEAREVFEDDEKTETFWNVYVDGDQCSSNPCHYRGTCKDGIGSYTCTCLSGYEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||    | | :|||:|  ||    || : |||| || | ::| :|:  | : ||:     |
Db        121 CERVLYKSCRVDNGNCWHFCKPVQNDIQCSCAEGYLLGEDGHSCVAGGNFSCGRNIKTRN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              || :          || :  :  :|  |  ::||           |:         ||| | :
Db        181 KREAS--------LPDFV--QSQNAILLKKSDNP---------SPD------IRIVNGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              || |||||||:|::|: : ||||||||  |:||||||: | |   | ||: |  ::  |
Db        216 CKLGECPWQAVLVDEKEDAFCGGTILSPIYVLTAAHCINQTKMISVVVGEINISRKNPGR 275

Qy        301 A-------VHEVEVVIK-----HNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDW 348
                    ||:  |  |     : :|   :||:|||:|::|||| |  || ||||  |:
Db        276 LLSVDKIYVHQKFVPPKKGYEFYEKFDLVSYDYDIAILQMKTPIQFSENVVPACLPTADF 335

Qy        349 AESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYD 408
              |   || |  |||||||  ||| ||  ||:|:||||||::| |||   ||  |||||||
Db        336 ANQVLMKQDFGIVSGFGRIFEKGPQSKTLKVLKVPYVDRHTCMLSSESPITPTMFCAGYD 395

Qy        409 TKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKT 468
              |  ||||||||||:| ::||:|:|||||||| ||: ||||:||||: |: || | |:
Db        396 TLPRDACQGDSGGPHITAYRDTHFITGIVSWGEGCAQTGKYGVYTKVSKFILWIKRIMRQ 455

Qy        469 RGLPKAKS 476
              : ||  :|
Db        456 K-LPSTES 462
```

RESULT 6
US-10-406-031-2
; Sequence 2, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
; APPLICANT: Masci, Paul Pantaleone
; APPLICANT: De Jersey, John
; APPLICANT: Lavin, Martin
; TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
; FILE REFERENCE: 15685-002001
; CURRENT APPLICATION NUMBER: US/10/406,031
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: AU 2003901033
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: AU PS1483
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Pseudonaja textilis
US-10-406-031-2

Query Match 47.0%; Score 1238; DB 12; Length 467;
Best Local Similarity 48.3%; Pred. No. 1.4e-94;
Matches 236; Conservative 74; Mismatches 139; Indels 40; Gaps 7;

```
Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              |   | | |:   |  |      ::|:: : ||  | |  |||| :|| | |::||||:||
Db          1 MAPQLLLCLILTFLWSLPEAESNVFLKSKVANRFLQRTKRANSLVEEFKSGNIERECIEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
               || ||||||||| :||  ||| | ||||| ::||  :| ||||:| |||||| |:||||
Db         61 RCSKEEAREVFEDDEKTETFWNVYVDGDQCSSNPCHYRGICKDGIGSYTCTCLSGYEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||    | | :|||:|  ||    || : |||| || | ::| :|:  | : ||:     |
Db        121 CERVLYKSCRVDNGNCWHFCKSVQNDIQCSCAEGYLLGEDGHSCVAGGNFSCGRNIKTRN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              || :          || :  : ::|  |  ::||           |:         ||| | :
Db        181 KREAS--------LPDFV--QSHNATLLKKSDNP---------SPD------IRIVNGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              || |||||||:|::|: : ||||||||  |:||||||: | |   | ||: |  ::  |
Db        216 CKLGECPWQAALVDDKKGVFCGGTILSPIYVLTAAHCINETETISVVVGEIDRSRAETGP 275

Qy        301 AVHEVEVVIKHNRFTKE-------------TYDFDIAVLRLKTPITFRMNVAPACLPERD 347
              :  |: |  | :|                  :||:|||::::|||| |  || |||||  |
Db        276 LL-SVDKVYVHKKFVPPKKSQEFYEKFDLVSYDYDIAIIQMKTPIQFSENVVPACLPTAD 334

Qy        348 WAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGY 407
              :|   || |  |||||||   |:|  |  ||:|:||||||::| |||:| ||  ||||||
Db        335 FANQVLMKQDFGIVSGFGGIFERGPNSKTLKVLKVPYVDRHTCMLSSNFPITPTMFCAGY 394

Qy        408 DTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467
              || :|||||||||||:| ::||:|:|||||||| |||||:||||||:: |: || | |:
Db        395 DTLPQDACQGDSGGPHITAYRDTHFITGIVSWGEGCARKGRYGIYTKLSKFIPWIKRIMR 454

Qy        468 TRGLPKAKS 476
                : ||  :|
Db        455 QK-LPSTES 462
```

RESULT 7
US-10-406-031-8
; Sequence 8, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
; APPLICANT: Masci, Paul Pantaleone
; APPLICANT: De Jersey, John
; APPLICANT: Lavin, Martin
; TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
; FILE REFERENCE: 15685-002001
; CURRENT APPLICATION NUMBER: US/10/406,031
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: AU 2003901033
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: AU PS1483
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Oxyuranus microlepidotus
US-10-406-031-8

Query Match 47.0%; Score 1238; DB 12; Length 467;
Best Local Similarity 48.3%; Pred. No. 1.4e-94;
Matches 236; Conservative 73; Mismatches 140; Indels 40; Gaps 7;

```
Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              |   | | |:   |  |      ::|:: : ||  | |  |||| :|| | |::||||:||
Db          1 MAPQLLLCLILTFLWSLPEAESNVFLKSKVANRFLQRTKRANSLFEEFRSGNIERECIEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
               || ||||||||| :||  ||| | ||||| ::||  :| ||||:| |||||| |:||||
Db         61 RCSKEEAREVFEDDEKTETFWNVYVDGDQCSSNPCHYRGTCKDGIGSYTCTCLFGYEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||    | | :|||:|  ||    || : |||| || | ::| :|:  | : ||:     |
Db        121 CERVLYKSCRVDNGNCWHFCKSVQNDIQCSCAEGYLLGEDGHSCVAGGNFSCGRNIKTRN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              || :          || :  : ::|  |  ::||           |:         ||| | :
Db        181 KREAS--------LPDFV--QSHNATLLKKSDNP---------SPD------IRIVNGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              || |||||||:|::|: : ||||||||  |:||||||: | |   | ||: |  ::  |
Db        216 CKLGECPWQAALVDDKKGVFCGGTILSPIYVLTAAHCINETETISVVVGEIDRSRAETGP 275

Qy        301 AVHEVEVVIKHNRFTKE-------------TYDFDIAVLRLKTPITFRMNVAPACLPERD 347
              :  |: |  | :|                  :||:|||::::|||| |  || |||||  |
Db        276 LL-SVDKVYVHKKFVPPKKSQEFYEKFDLVSYDYDIAIIQMKTPIQFSENVVPACLPTAD 334

Qy        348 WAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGY 407
              :|   || |  |||||||   |:|  |  ||:|:||||||::| |||:| ||  ||||||
Db        335 FANQVLMKQDFGIVSGFGGIFERGPNSKTLKVLKVPYVDRHTCMLSSNFPITPTMFCAGY 394

Qy        408 DTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467
              || :|||||||||||:| ::||:|:|||||||| |||||:||||||:: |: || | |:
Db        395 DTLPQDACQGDSGGPHITAYRDTHFITGIVSWGEGCARKGRYGIYTKLSKFIPWIKRIMR 454

Qy        468 TRGLPKAKS 476
                : ||  :|
Db        455 QK-LPSTES 462
```

```
Db        121 CERVLYKSCRVDNGNCWHFCKPVQNDIQCSCAEGYLLGEDGHSCVAGGNFSCGRNIKTRN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Db        181 KREAS--------LPDFV--QSQNATLLKKSDNP---------SPD------IRIVNGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Db        216 CKLGECPWQAVLVDEKEGVFCGGTILSPIYVLTAAHCINQTEKISVVVGEIDKSRVETGH 275

Qy        301 AVHEVEVVIKHNRFTKE------------TYDFDIAVLRLKTPITFRMNVAPACLPERD 347

Db        276 LL-SVDKIYVHKKFVPPKKGYKFYEKFDLVSYDYDIAIIQMKTPIQFSENVVPACLPTAD 334

Qy        348 WAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGY 407

Db        335 FANQVLMKQDFGIISGFGRIFEKGPKSNTLKVLKVPYVDRHTCMVSSESPITPTMFCAGY 394

Qy        408 DTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467

Db        395 DTLPRDACQGDSGGPHITAYRDTHFITGIVSWGEGCAKKGKYGIYTKVSKFILWIKRIMR 454

Qy        468 TRGLPKAKS 476

Db        455 QK-LPSTES 462


RESULT 8
US-10-406-031-17
; Sequence 17, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
;  APPLICANT: Masci, Paul Pantaleone
;  APPLICANT:  De Jersey, John
;  APPLICANT:  Lavin, Martin
;  TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
;  FILE REFERENCE: 15685-002001
;  CURRENT APPLICATION NUMBER: US/10/406,031
;  CURRENT FILING DATE:  2003-04-02
;  PRIOR APPLICATION NUMBER: AU 2003901033
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: AU PS1483
;  PRIOR FILING DATE: 2002-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 455
;   TYPE: PRT
;   ORGANISM: Tropidechis carinatus
US-10-406-031-17

  Query Match           46.4%;  Score 1223;  DB 12;  Length 455;
  Best Local Similarity   48.5%;  Pred. No. 2.5e-93;
  Matches  233;  Conservative   70;  Mismatches  141;  Indels   36;  Gaps    6;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Db          1 MAPQLLLCLILTFLWSLPEAESNVFLKSKVANRFLQRTKRSNSLFEEIRPGNIERECIEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Db         61 KCSKEEAREVFEDNEKTETFWNVYVDGDQCSSNPCHYRGTCKDGIGSYTCTCLPNYEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Db        121 CEKVLYQSCRVDNGNCWHFCKRVQSETQCSCAESYRLGVDGHSCVAEGDFSCGRNIKARN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Db        181 KREAS--------LPDFV--QSQKATLLKKSDNP---------SPD------IRIVNGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Db        216 CKLGECPWQAVLINEKGEVFCGGTILSPIHVLTAAHCINQTKSVSVIVGEIDISRKETRR 275

Qy        301 --AVHEVEVVIK---------HNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWA 349

Db        276 LLSVDKIYVHTKFVPPNYYYVHQNFDRVAYDYDIAIIRMKTPIQFSENVVPACLPTADFA 335

Qy        350 ESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDT 409

Db        336 NEVLMKQDSGIVSGFGRIQFKQPTSNTLKVITVPYVDRHTCMLSSDFRITQNMFCAGYDT 395

Qy        410 KQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTR 469

Db        396 LPQDACQGDSGGPHITAYRDTHFITGIISWGEGCARKGKYGVYTKVSKFIPWIKKIMSLK 455


RESULT 9
US-10-406-031-11
; Sequence 11, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
;  APPLICANT: Masci, Paul Pantaleone
;  APPLICANT:  De Jersey, John
;  APPLICANT:  Lavin, Martin
;  TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
;  FILE REFERENCE: 15685-002001
;  CURRENT APPLICATION NUMBER: US/10/406,031
;  CURRENT FILING DATE:  2003-04-02
;  PRIOR APPLICATION NUMBER: AU 2003901033
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: AU PS1483
;  PRIOR FILING DATE: 2002-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 454
;   TYPE: PRT
;   ORGANISM: Pseudechis porphyriacus
US-10-406-031-11

  Query Match           46.4%;  Score 1222;  DB 12;  Length 454;
  Best Local Similarity   47.0%;  Pred. No. 3e-93;
  Matches  228;  Conservative   75;  Mismatches  148;  Indels   34;  Gaps    6;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60

Db          1 MAPQLLLCLILTFLWSLPEAESNVFLKSKEANRFLQRTKRSNSLFEEFRPGNIERECIEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120

Db         61 KCSKEEAREIFKDNEKTEAFWNVYVDGDQCSSNPCHYGGTCKDGIGSYTCTCLPNYEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180

Db        121 CEHLLFKSCRFFNGNCWHFCKPVQNDTQCSCAESYRLGDDGHSCVAEGDFSCGRNIKARN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240

Db        181 KREAS--------LPDFV--QSQNATLLKKSDNP---------SPD------IRIINGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300

Db        216 CKLGECPWQAVLLDKEGDVFCGGTILSPIYVLTAAHCITQSKHISVVVGEIDISRKETRH 275

Qy        301 AVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGI 360

Db        276 LL-SVDKAYVHTKFVLATYDYDIAIIQLKTPIQFSENVVPACLPTADFANQVLMKQDFGI 334

Qy        361 VSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSG 420

Db        335 ISGFGHTRSGGQTSNTLKVVTIPYVDRHTCMLSSDFRITPNMFCAGYDTLPRDACQGDSG 394

Qy        421 GPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGLPKAKSHAPE 480
```

```
Db        395 GPHITAYRDTHFITGIISWGEGCAKKGKYGVYTKVSNFIPWIKAVM--------RKHQPS 446

Qy        481 VITSS 485
              :|:
Db        447 TESST 451


RESULT 10
US-10-406-031-14
; Sequence 14, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
;  APPLICANT: Masci, Paul Pantaleone
;  APPLICANT:  De Jersey, John
;  APPLICANT:  Lavin, Martin
;  TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
;  FILE REFERENCE: 15685-002001
;  CURRENT APPLICATION NUMBER: US/10/406,031
;  CURRENT FILING DATE:  2003-04-02
;  PRIOR APPLICATION NUMBER: AU 2003901033
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: AU PS1483
;  PRIOR FILING DATE: 2002-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 453
;   TYPE: PRT
;   ORGANISM: No. US20040043017A1echis scutatus
US-10-406-031-14

  Query Match             46.2%;  Score 1216;  DB 12;  Length 453;
  Best Local Similarity   48.6%;  Pred. No. 9.5e-93;
  Matches  233;  Conservative   69;  Mismatches  141;  Indels   36;  Gaps    6;

Qy          1 MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEE 60
              |   | | |:   |  |     ::|:: : ||  | |  |:|| ||:: |::||||:|| 
Db          1 MAPQLLLCLILTFLWSLPEAESNVFLKSKVANRFLQRTKRSNSLFEEIRPGNIERECIEE 60

Qy         61 TCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN 120
               || |||||||||::||  ||| | ||||| ::||  :| ||||:| ||||||  :|||| 
Db         61 KCSKEEAREVFEDNEKTETFWNVYVDGDQCSSNPCHYRGTCKDGIGSYTCTCLPNYEGKN 120

Qy        121 CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERR 180
              ||    | |   ||:| ||   |:   |||| | | :| :|:  | : ||:    |
Db        121 CEKVLFKSCRAFNGNCWHFCKRVQSETQCSCAESYLLGVDGHSCVAEGDFSCGRNIKARN 180

Qy        181 KRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQE 240
              ||  :        || :  :   |  |  ::||          |:       ||| | :
Db        181 KREAS--------LPDFV--QSQKATVLKKSDNP---------SPD------IRIVNGMD 215

Qy        241 CKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGE 300
              || |||||||:||||: | ||||||||  ::|||||: | |    | ||: :  ::|
Db        216 CKLGECPWQAVLINEKGEVFCGGTILSPIHVLTAAHCINQTKSVSVIVGEIDISRKETRR 275

Qy        301 AVHEVEVVIKHNRFT----------KETYDFDIAVLRLKTPITFRMNVAPACLPERDWAE 350
                :  |: :  | :|           :  ||:|||::|:|||| |  || |||||  |:|: 
Db        276 LL-SVDKIYVHKKFVPPNSYYQNIDRFAYDYDIAIIRMKTPIQFSENVVPACLPTADFAK 334

Qy        351 STLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTK 410
                || | :|||||||||   |  |  ||:: |||||::| |||:| ||||||||||||
Db        335 EVLMKQDSGIVSGFGRTQSIGYTSNILKVITVPYVDRHTCMLSSNFRITQNMFCAGYDTL 394

Qy        411 QEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTR 469
              :||||||||||:| : ||:|||||:|||| |||||||:||||: |: || : |  :
Db        395 PQDACQGDSGGPHITAYGDTHFVTGIISWGEGCARKGKYGVYTKVSNFIPWIKKIMSLK 453


RESULT 11
US-10-406-031-31
; Sequence 31, Application US/10406031
; Publication No. US20040043017A1
; GENERAL INFORMATION:
;  APPLICANT: Masci, Paul Pantaleone
;  APPLICANT:  De Jersey, John
;  APPLICANT:  Lavin, Martin
;  TITLE OF INVENTION: PROTHROMBIN ACTIVATING PROTEIN
;  FILE REFERENCE: 15685-002001
;  CURRENT APPLICATION NUMBER: US/10/406,031
;  CURRENT FILING DATE:  2003-04-02
;  PRIOR APPLICATION NUMBER: AU 2003901033
;  PRIOR FILING DATE: 2003-03-07
;  PRIOR APPLICATION NUMBER: AU PS1483
;  PRIOR FILING DATE: 2002-04-03
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
;   LENGTH: 376
;   TYPE: PRT
;   ORGANISM: Tropidechis carinatus
US-10-406-031-31

  Query Match             43.1%;  Score 1135.5;  DB 12;  Length 376;
  Best Local Similarity   47.7%;  Pred. No. 3.8e-86;
  Matches  209;  Conservative   55;  Mismatches  103;  Indels   71;  Gaps    3;

Qy         41 ANSFLEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGK 100
              :||  ||:: |::||||:|| || |||||||||::||  ||| | ||||| ::||  :|
Db          1 SNSLFEEIRPGNIERECIEEKCSKEEAREVFEDNEKTETFWNVYVDGDQCSSNPCHYRGT 60

Qy        101 CKDGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN 160
              ||||:| ||||||  :||||||    : | :|||:|  ||    |:   |||| | | :
Db         61 CKDGIGSYTCTCLPNYEGKNCEKVLYQSCRVDNGNCWHFCKRVQSETQCSCAESYRLGVD 120

Qy        161 GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDF 220
              | :|:  | : ||:     | |
Db        121 GHSCVAEGDFSCGRNIKARNK---------------------------------------- 141

Qy        221 NQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQ 280
                          || | :|| |||||||:||||: | ||||||||  ::|||||: |
Db        142 --------------IVNGMDCKLGECPWQAVLINEKGEVFCGGTILSPIHVLTAAHCINQ 187

Qy        281 AKRFKVRVGDRNTEQEEGGEAVHEVEVVIK---------HNRFTKETYDFDIAVLRLKTP 331
              | |          :     :| :: |  |          |  | :  ||:|||::|:|||
Db        188 TKSVK---------ETRRLLSVDKIYVHTKFVPPNYYYVHQNFDRVAYDYDIAIIRMKTP 238

Qy        332 ITFRMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCK 391
              | |  || |||||  |:|    || | :||||||||    |     |  ||:: ||||||::|
Db        239 IQFSENVVPACLPTADFANEVLMKQDSGIVSGFGRIQFKQPTSNTLKVITVPYVDRHTCM 298

Qy        392 LSSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGI 451
              ||| | ||||||||||||  :||||||||||:| ::||:|:|||:|||| ||||||||:
Db        299 LSSDFRITQNMFCAGYDTLPQDACQGDSGGPHITAYRDTHFITGIISWGEGCARKGKYGV 358

Qy        452 YTKVTAFLKWIDRSMKTR 469
              ||||: |: || : |  :
Db        359 YTKVSKFIPWIKKIMSLK 376


RESULT 12
US-09-884-901-3
; Sequence 3, Application US/09884901
; Patent No. US20020076798A1
; GENERAL INFORMATION:
;  APPLICANT: Miao, Carol
;  APPLICANT:  Kay, Mark
;  TITLE OF INVENTION: Liver-Specific Gene Expression Cassettes, and Methods of Use
;  FILE REFERENCE: UOFW-1-17396
;  CURRENT APPLICATION NUMBER: US/09/884,901
;  CURRENT FILING DATE:  2001-06-18
;  PRIOR APPLICATION NUMBER: US 60/212,902
;  PRIOR FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 461
```

```
;   TYPE: PRT
;   ORGANISM: HomoSapien
US-09-884-901-3

  Query Match           39.5%;  Score 1041.5;  DB 9;  Length 461;
  Best Local Similarity   44.8%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   76;  Mismatches  157;  Indels   23;  Gaps   10;

Qy          9 LLSASLAGLLLLGE-SLFIRREQANNILARVTRANS-FLEEMKKGHLERECMEETCSYEE 66
              |::  | | ||  | ::|:  | || || |  | ||  ||| :|:|||||||| ||:|| 
Db         14 LITICLLGYLLSAECTVFLDHENANKILNRPKRYNSGKLEEFVQGNLERECMEEKCSFEE 73

Qy         67 AREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELFTR 126
              ||||||::::| ||| :| ||||||::|| | | ||| :  | | |  |||||||||
Db         74 AREVFENTERTTEFWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELDV- 132

Qy        127 KLCSLDNGDCDQFC-HEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSVA 185
                |:: || |:||| :   | |||||  || ||:| |:| |  |:|||: :: :     :
Db        133 -TCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQ----TS 187

Qy        186 QATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECKDGE 245
              : | :   ||       | :   |   |  || :  |: ||:|||:: | |:
Db        188 KLTRAEAVFPD------VDYVNSTEAETILD----NITQSTQSFNDFTRVVGGEDAKPGQ 237

Qy        246 CPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAVHEV 305
              ||| :| | : : |||||:|::| :|:|||||:    :  |  |: | |: |  |    |
Db        238 FPWQVVL-NGKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNV 296

Qy        306 EVVIKHNRFTK--ETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVSG 363
              :| |: :      |: |||:| |  |:      | | |: :::: :  :   :| |||
Db        297 IRIIPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYT-NIFLKFGSGYVSG 355

Qy        364 FGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGPH 423
              :||   ||| :  |: | || ||| :|  |: | |  ||||||:     |:|||||||
Db        356 WGRVFHKGRSALVLQYLRVPLVDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPH 415

Qy        424 VTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467
              ||  : | |:|||:|||| || ||||||||||: :: ||   |
Db        416 VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTK 459


RESULT 13
US-10-234-406-8
; Sequence 8, Application US/10234406
; Publication No. US20030109478A1
; GENERAL INFORMATION:
;  APPLICANT: FEWEL, Jason G.
;  APPLICANT:  MACLAUGHLIN, Fiona
;  APPLICANT:  SMITH, Louis C.
;  APPLICANT:  NICOL, Francois
;  APPLICANT:  ROLLAND, Alain
;  TITLE OF INVENTION: NUCLEIC ACID FORMULATIONS FOR GENE DELIVERY AND METHODS OF USE
;  FILE REFERENCE: 54964.8303.US01
;  CURRENT APPLICATION NUMBER: US/10/234,406
;  CURRENT FILING DATE:  2002-09-03
;  PRIOR APPLICATION NUMBER: US 60/187,236
;  PRIOR FILING DATE: 2000-03-03
;  PRIOR APPLICATION NUMBER: US 60/261,751
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: PCT/US01/06953
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Expression plasmid pFN1645 having codon optimized sequence encodi
;   OTHER INFORMATION: ng for human coagulation factor IX (786) ... (2171).
US-10-234-406-8

  Query Match           39.5%;  Score 1041.5;  DB 14;  Length 461;
  Best Local Similarity   44.8%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   76;  Mismatches  157;  Indels   23;  Gaps   10;

Qy          9 LLSASLAGLLLLGE-SLFIRREQANNILARVTRANS-FLEEMKKGHLERECMEETCSYEE 66
              |::  | | ||  | ::|:  | || || |  | ||  ||| :|:|||||||| ||:|| 
Db         14 LITICLLGYLLSAECTVFLDHENANKILNRPKRYNSGKLEEFVQGNLERECMEEKCSFEE 73

Qy         67 AREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELFTR 126
              ||||||::::| ||| :| ||||||::|| | | ||| :  | | |  |||||||||
Db         74 AREVFENTERTTEFWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELDV- 132

Qy        127 KLCSLDNGDCDQFC-HEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSVA 185
                |:: || |:||| :   | |||||  || ||:| |:| |  |:|||: :: :     :
Db        133 -TCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQ----TS 187

Qy        186 QATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECKDGE 245
              : | :   ||       | :   |   |  || :  |: ||:|||:: | |:
Db        188 KLTRAEAVFPD------VDYVNSTEAETILD----NITQSTQSFNDFTRVVGGEDAKPGQ 237

Qy        246 CPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAVHEV 305
              ||| :| | : : |||||:|::| :|:|||||:    :  |  |: | |: |  |    |
Db        238 FPWQVVL-NGKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNV 296

Qy        306 EVVIKHNRFTK--ETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVSG 363
              :| |: :      |: |||:| |  |:      | | |: :::: :  :   :| |||
Db        297 IRIIPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYT-NIFLKFGSGYVSG 355

Qy        364 FGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGPH 423
              :||   ||| :  |: | || ||| :|  |: | |  ||||||:     |:|||||||
Db        356 WGRVFHKGRSALVLQYLRVPLVDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPH 415

Qy        424 VTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467
              ||  : | |:|||:|||| || ||||||||||: :: ||   |
Db        416 VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTK 459


RESULT 14
US-10-038-854-92
; Sequence 92, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
;  APPLICANT: Spytek, Kimberly A
;  APPLICANT:  Li, Li
;  APPLICANT:  Wolenc, Adam R
;  APPLICANT:  Vernet, Corine
;  APPLICANT:  Eisen, Andrew J
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Tchernev, Velizar
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gusev, Vladimir Y
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Guo, Xiaojia S
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  MacDougall, John R
;  TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-230
;  CURRENT APPLICATION NUMBER: US/10/038,854
;  CURRENT FILING DATE:  2003-01-22
;  PRIOR APPLICATION NUMBER: 60/258,928
```

```
;  PRIOR FILING DATE: 2000-12-29
;  PRIOR APPLICATION NUMBER: 60/259,415
;  PRIOR FILING DATE: 2001-01-02
;  PRIOR APPLICATION NUMBER: 60/259,785
;  PRIOR FILING DATE: 2001-01-04
;  PRIOR APPLICATION NUMBER: 60/269,814
;  PRIOR FILING DATE: 2001-02-20
;  PRIOR APPLICATION NUMBER: 60/279,832
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/279,833
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/279,863
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/283,889
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: 60/284,447
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/286,683
;  PRIOR FILING DATE: 2001-04-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 411
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;    LENGTH: 461
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-038-854-92

  Query Match             39.5%;  Score 1041.5;  DB 16;  Length 461;
  Best Local Similarity   44.8%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   76;  Mismatches  157;  Indels   23;  Gaps   10;

Qy          9 LLSASLAGLLLLGE-SLFIRREQANNILARVTRANS-FLEEMKKGHLERECMEETCSYEE 66
              |::  | | ||  | ::|:  | || || |  | ||  |||  :|:||||||||| ||:||
Db         14 LITICLLGYLLSAECTVFLDHENANKILNRPKRYNSGKLEEFVQGNLERECMEEKCSFEE 73

Qy         67 AREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELFTR 126
              ||||||::::| ||| :| |||||||::|| | | ||| :  | | |  |||||||||
Db         74 AREVFENTERTTEFWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELDV- 132

Qy        127 KLCSLDNGDCDQFC-HEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSVA 185
                |:: || |:||| :   | |||||  || ||:| |:| |  |:|||: :: :    :
Db        133 -TCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQ----TS 187

Qy        186 QATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECKDGE 245
              : | :    ||      | :    |  |   |||  :  |: ||:|||:: | |:
Db        188 KLTRAEAVFPD------VDYVNSTEAETILD----NITQSTQSFNDFTRVVGGEDAKPGQ 237

Qy        246 CPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAVHEV 305
              ||| :| | : : ||||:|::| :|:|||||:    :  |  |: | |: |  |     |
Db        238 FPWQVVL-NGKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNV 296

Qy        306 EVVIKHNRFTK--ETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVSG 363
              :| |: :      |: |||:|  |:     | | |: :::: :  :  :| |||
Db        297 IRIIPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYT-NIFLKFGSGYVSG 355

Qy        364 FGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGPH 423
              :||   ||| :  |: | || ||| :|  |: | |  ||||||:     |:|||||||||
Db        356 WGRVFHKGRSALVLQYLRVPLVDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPH 415

Qy        424 VTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467
              ||  : | |:|||:|||| || ||||||||||: :: ||    |
Db        416 VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTK 459


RESULT 15
US-10-239-498A-5
; Sequence 5, Application US/10239498A
; Publication No. US20040023333A1
; GENERAL INFORMATION:
;  APPLICANT: Hauser, Charlotte
;  APPLICANT:  Horster, Andrea
;  APPLICANT:  Schroder, Carola
;  APPLICANT:  Lehnerer, Michael
;  TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
;  TITLE OF INVENTION:  Human Cell Lines
;  FILE REFERENCE: 80977.0001
;  CURRENT APPLICATION NUMBER: US/10/239,498A
;  CURRENT FILING DATE:  2003-07-08
;  PRIOR APPLICATION NUMBER: PCT/EP01/03220
;  PRIOR FILING DATE: 2001-03-21
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;    LENGTH: 461
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-10-239-498A-5

  Query Match             39.5%;  Score 1041.5;  DB 16;  Length 461;
  Best Local Similarity   44.8%;  Pred. No. 3.3e-78;
  Matches  208;  Conservative   76;  Mismatches  157;  Indels   23;  Gaps   10;

Qy          9 LLSASLAGLLLLGE-SLFIRREQANNILARVTRANS-FLEEMKKGHLERECMEETCSYEE 66
              |::  | | ||  | ::|:  | || || |  | ||  |||  :|:||||||||| ||:||
Db         14 LITICLLGYLLSAECTVFLDHENANKILNRPKRYNSGKLEEFVQGNLERECMEEKCSFEE 73

Qy         67 AREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKNCELFTR 126
              ||||||::::| ||| :| |||||||::|| | | ||| :  | | |  |||||||||
Db         74 AREVFENTERTTEFWKQYVDGDQCESNPCLNGGSCKDDINSYECWCPFGFEGKNCELDV- 132

Qy        127 KLCSLDNGDCDQFC-HEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLERRKRSVA 185
                |:: || |:||| :   | |||||  || ||:| |:| |  |:|||: :: :    :
Db        133 -TCNIKNGRCEQFCKNSADNKVVCSCTEGYRLAENQKSCEPAVPFPCGRVSVSQ----TS 187

Qy        186 QATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIVGGQECKDGE 245
              : | :    ||      | :    |  |   |||  :  |: ||:|||:: | |:
Db        188 KLTRAEAVFPD------VDYVNSTEAETILD----NITQSTQSFNDFTRVVGGEDAKPGQ 237

Qy        246 CPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRVGDRNTEQEEGGEAVHEV 305
              ||| :| | : : ||||:|::| :|:|||||:    :  |  |: | |: |  |     |
Db        238 FPWQVVL-NGKVDAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNV 296

Qy        306 EVVIKHNRFTK--ETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMTQKTGIVSG 363
              :| |: :      |: |||:|  |:     | | |: :::: :  :  :| |||
Db        297 IRIIPHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYT-NIFLKFGSGYVSG 355

Qy        364 FGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQEDACQGDSGGPH 423
              :||   ||| :  |: | || ||| :|  |: | |  ||||||:     |:|||||||||
Db        356 WGRVFHKGRSALVLQYLRVPLVDRATCLRSTKFTIYNNMFCAGFHEGGRDSCQGDSGGPH 415

Qy        424 VTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMK 467
              ||  : | |:|||:|||| || ||||||||||: :: ||    |
Db        416 VTEVEGTSFLTGIISWGEECAMKGKYGIYTKVSRYVNWIKEKTK 459


Search completed: April 14, 2004, 15:46:09
Job time : 48 secs
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "hope1-ags" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "new.key":
  hope1 (AA) ID   hope1 AA preliminary pattern
1       followed by
2         e
2         q or s or h or y or e
2         s or q or i or t or n or p
2         f or t or s or p or l or i
2         n or s or k or m or t or p
2         d or k or t or e
2         f or l or r or i
2         t or s or n
2         r
2         i or v or a

Selected data banks and files:

   Data bank : A-GeneSeq 35.2, all entries

                      -- Output Parameters --

Format Options:                          File Options:
  Nucleic acid code matching   Exact       Indirect file                No
  Find non-matching hits only  No          Sequence or key file         No
  Report key used              Yes         List of hits                 Yes
  Note position of hit         Yes         Hit display                  Yes
  Display full annotations     Yes         Name and annotations         Yes
  Sequence context             10


                      -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 R30729 ; p100 protein from human herpes virus type 6.
  (from "A-GeneSeq 35.2")
 ID   R30729 standard; Protein; 870 AA.
 AC   R30729;
 DT   20-MAY-1993  (first entry)
 DE   p100 protein from human herpes virus type 6.
 KW   antibodys; monoclonal antibody; ELISA assays; CMV; cytomegalovirus.
 OS   Human herpes virus type 6.
 PN   EP-524421-A.
 PD   27-JAN-1993.
 PF   15-JUN-1992; 110047.
 PR   08-JUL-1991; EP-111338.
 PA   (BEHW ) BEHRINGWERKE AG.
 PI   Fleckenstein B,  Neipel F;
 DR   WPI; 93-028531/04.
 DR   P-PSDB; R30729.
 PT   Human herpes virus type 6 protein p100 DNA sequence - useful in
 PT   prophylaxis, treatment and differential diagnosis of human herpes
 PT   virus-6 infections
 PS   Claim 1; Page 12; 25pp; English.
 CC   This sequence is the p100 protein from human herpes virus type6.
 CC   The protein and antibodies to it can be used for treatment or
 CC   prevention of HHV-6 infections. The DNA , protein and Ab are also
 CC   useful in eg. ELISA assays esp. for differentiating between HHV-6
 CC   and cytomegalovirus infections. These assays are more sensitive
 CC   and specific than immunofluorescence methods currently used.
 SQ   Sequence   870 AA;
 SQ   43 A; 42 R; 67 N; 69 D; 0  B; 1  C; 33 Q; 57 E; 0  Z; 54 G; 16 H;
 SQ   42 I; 94 L; 63 K; 17 M; 33 F; 35 P; 86 S; 46 T; 8  W; 15 Y; 49 V;
 Found using 'hope1' (new.key)

...

                  |--------|
110     enmftnkekfesqfsdinrallrlgnfikw
                  120      129

...

-------
1 match found in sequence:
 R42456 ; Enzyme involved in eicosapentaenoic acid (EPA) synthesis.
  (from "A-GeneSeq 35.2")
 ID   R42456 standard; Protein; 543 AA.
 AC   R42456;
 DT   27-MAY-1994  (first entry)
 DE   Enzyme involved in eicosapentaenoic acid (EPA) synthesis.
 KW   EPA; eicosapentaenoic acid synthetase; drug; anticoagulant;
 KW   hypolipemic; hypoglycemic; antihypertensive; anticancer; pesticide;
 KW   foodstuff; additive.
 OS   Shewanella putrefaciens.
 PN   WO9323545-A.
 PD   25-NOV-1993.
 PF   14-MAY-1993; J00641.
 PR   15-MAY-1992; JP-147945.
 PA   (SAGA ) SAGAMI CHEM RES CENTRE.
 PI   Kato S,  Kondo K,  Yamada A,  Yazawa K;
 DR   WPI; 93-386577/48.
 DR   N-PSDB; Q51128.
 PT   Gene coding for eicosa-penta:enoic acid synthetase - is isolated
 PT   from Pseudomonas, Alteromonas or Shewanella and used for
 PT   recombinant prodn. of eicosa-penta:enoic acid
 PS   Claim 10; Page 91-94; 106pp; Japanese.
 CC   EPA is useful as a drug, having anticoagulant, hypolipemic,
 CC   hypoglycemic, antihypertensive and anticancer activity.  It is also
 CC   a pesticide and is useful as a nutritional foodstuff and animal feed
 CC   additive.
 SQ   Sequence   543 AA;
 SQ   69 A; 28 R; 25 N; 26 D; 0  B; 4  C; 24 Q; 33 E; 0  Z; 41 G; 9  H;
 SQ   27 I; 46 L; 25 K; 20 M; 17 F; 26 P; 30 S; 28 T; 8  W; 19 Y; 38 V;
 Found using 'hope1' (new.key)

...

                  |--------|
20      snisfdvqvmeqqlkdfsracyvvnhadhg
                  30       39

...

-------
1 match found in sequence:
 R99465 ; Biosynthetic enzyme of icosapentaenoic acid synthase.
  (from "A-GeneSeq 35.2")
 ID   R99465 standard; Protein; 543 AA.
 AC   R99465;
 DT   30-JAN-1997  (first entry)
 DE   Biosynthetic enzyme of icosapentaenoic acid synthase.
 KW   Icosapentaenoic acid synthase; EPA; drugs; agrochemicals;
 KW   foodstuffs; animal feed; lipid balance correction; antihypertensive;
 KW   antiinflammatory; anticancer agent.
 OS   Shewanella putrefaciens.
 PN   WO9621735-A1.
 PD   18-JUL-1996.
 PF   12-JAN-1996; J00030.
 PR   13-JAN-1995; JP-004299.
 PA   (SAGA ) SAGAMI CHEM RES CENTRE.
 PI   Kato S,  Kondo K,  Yamada A,  Yazawa K;
```

DR WPI; 96-342288/34.
DR N-PSDB; T34137.
PT Production of icosapentaenoic acid using transformed E. coli - uses
PT DNA coding for icosapentaenoic acid synthase derived from Shewanella
PT strain
PS Claim 7; Page 128-131; 145pp; English.
CC The DNA sequence (T34137) which encodes the biosynthetic enzymes of
CC icosapentaenoic acid (EPA) can be used to transform Escherichia coli.
CC The DNA sequence allows efficient microbial production of EPA, which
CC is a raw material for drugs, agrochemicals, foods and animal
CC feedstuffs. EPA is also useful for lipid balance correction and as
CC an antihypertensive, antiinflammatory and anticancer agent.
SQ Sequence 543 AA;
SQ 69 A; 28 R; 25 N; 26 D; 0 B; 4 C; 24 Q; 33 E; 0 Z; 41 G; 9 H;
SQ 27 I; 46 L; 25 K; 20 M; 17 F; 26 P; 30 S; 28 T; 8 W; 19 Y; 38 V;
Found using 'hope1' (new.key)

...

|--------|
20 snisfdvqvmeqqlkdfsracyvvnhadhg
30 39

...

---
1 match found in sequence:
W37053 ; S. putrefaciens EPO biosynthesis gene cluster ORF9 product.
(from "A-GeneSeq 35.2")
ID W37053 standard; Protein; 543 AA.
AC W37053;
DT 03-JUL-1998 (first entry)
DE S. putrefaciens EPO biosynthesis gene cluster ORF9 product.
KW SCRC-2874; FERM BP-1625; eicosapentaenoic acid; EPA;
KW biosynthesis gene cluster; synthetase.
OS Shewanella putrefaciens.
PN WO9801565-A1.
PD 15-JAN-1998.
PF 09-JUL-1997; J02371.
PR 10-JUL-1996; JP-180845.
PA (SAGA) SAGAMI CHEM RES CENTRE.
PI Kato S, Kondo K, Yamada A, Yazawa K;
DR WPI; 98-101060/09.
DR N-PSDB; V00503.
PT Eicosapentaenoic acid produced by culture of transformed Escherichia
PT coli - containing an eicosapentaenoic acid synthetase gene derived
PT from the marine microorganism Shewanella
PS Example 1; Pages 93-97; 110pp; Japanese.
CC The present sequence is encoded by the Shewanella putrefaciens
CC SCRC-2874 (FERM BP-1625) eicosapentaenoic acid (EPA) biosynthesis
CC gene cluster.
CC A novel EPA (useful in drugs, pesticides, foods and feedstuffs) is
CC encoded by synthetase enzyme gene sequences comprising parts of the
CC full sequence of the synthetase gene from the marine microorganism
CC S. putrefaciens SCRC-2874 (FERM BP-1625), in which at least 1 of
CC the 9 open reading frames (ORF) (numbered 2-10) in the gene have
CC been deleted. In particular the gene sequences comprising the
CC following parts of the full gene:
CC (1) bases 8081-9441, 12314-13084 and 13889-32520;
CC (2) bases 8081-9441, 12314-13084, 13889-32520 and 34627-35559;
CC (3) bases 8081-9441, 12314-13084 and 13889-35559;
CC (4) bases 8081-9441, 9681-13084 and 13889-32520;
CC (5) bases 8081-9441, 9681-13084, 13889-32520 and 34627-35564; and
CC (6) bases 8081-9441, 9681-13084 and 13889-35564, are claimed.
SQ Sequence 543 AA;
SQ 69 A; 28 R; 25 N; 26 D; 0 B; 4 C; 24 Q; 33 E; 0 Z; 41 G; 9 H;
SQ 27 I; 46 L; 25 K; 20 M; 17 F; 26 P; 30 S; 28 T; 8 W; 19 Y; 38 V;
Found using 'hope1' (new.key)

...

|--------|
20 snisfdvqvmeqqlkdfsracyvvnhadhg
30 39

...

-- Search Statistics --

Times:	CPU	Total Elapsed
	00:00:20.03	00:00:23.00

Number of sequences searched: 170751
Number of sequence hits: 4
Number of separate matches: 4
Number of sequence hits saved: 0

A-generic

```
! FINDPATTERNS on geneseqp:* allowing 0 mismatches
!       1 E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R,I)(T,S,N)R

(2) AAR30729 ck: 500  len: 870  ! Aar30729 p100 protein from human herpes vir
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(S)(Q)(F)(S)(D)(I)(N)R(A)
      120: NKEKF  ESQFSDINRA  LLRLG

(7) AAR42456 ck: 3609  len: 543  ! Aar42456 Enzyme involved in eicosapentaenoi
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(Q)(Q)(L)(K)(D)(F)(S)R(A)
       30: DVQVM  EQQLKDFSRA  CYVVN

(8) AAR99465 ck: 3609  len: 543  ! Aar99465 Biosynthetic enzyme of icosapentae
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(Q)(Q)(L)(K)(D)(F)(S)R(A)
       30: DVQVM  EQQLKDFSRA  CYVVN

(11) AAW37053 ck: 3609  len: 543  ! Aaw37053 S. putrefaciens EPO biosynthesis g
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(Q)(Q)(L)(K)(D)(F)(S)R(A)
       30: DVQVM  EQQLKDFSRA  CYVVN

(12) AAW89403 ck: 3609  len: 543  ! Aaw89403 S. putrefaciens PKS-like cluster C
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(Q)(Q)(L)(K)(D)(F)(S)R(A)
       30: DVQVM  EQQLKDFSRA  CYVVN

(1) AAB10470 ck: 3609  len: 543  ! Aab10470 Shewanella putrefaciens PKS protei
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(Q)(Q)(L)(K)(D)(F)(S)R(A)
       30: DVQVM  EQQLKDFSRA  CYVVN

(6) AAG52216 ck: 7182  len: 274  ! Aag52216 Arabidopsis thaliana protein fragm
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(E)(N)(S)(N)(T)(L)(N)R(V)
      168: QQVNP  EENSNTLNRV  NLGEQ

(3) AAG52217 ck: 376  len: 217  ! Aag52217 Arabidopsis thaliana protein fragm
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(E)(N)(S)(N)(T)(L)(N)R(V)
      111: QQVNP  EENSNTLNRV  NLGEQ

(4) AAG52218 ck: 9159  len: 151  ! Aag52218 Arabidopsis thaliana protein fragm
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(E)(N)(S)(N)(T)(L)(N)R(V)
       45: QQVNP  EENSNTLNRV  NLGEQ

(5) AAG64458 ck: 9499  len: 542  ! Aag64458 S. putrefaciens eicosapentanoic ac
                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                  E(Q)(Q)(L)(K)(D)(F)(S)R(A)
       30: DVQVM  EQQLKDFSRA  CYVVN

(9) AAU65346 ck: 8069  len: 79  ! Aau65346 Propionibacterium acnes immunogen
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,
                  E(S)(P)(P)(P)(T)(R)(S)R(I)
       40: RPAAK  ESPPPTRSRI  SRLGR

(10) AAU66759 ck: 99  len: 113  ! Aau66759 Propionibacterium acnes immunogen
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,
                  E(S)(P)(P)(P)(T)(R)(S)R(I)
       74: RPAAK  ESPPPTRSRI  SRLGR

(14) ABM63278 ck: 99  len: 113  ! Abm63278 Propionibacterium acnes predicted
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,
                  E(S)(P)(P)(P)(T)(R)(S)R(I)
       74: RPAAK  ESPPPTRSRI  SRLGR

(13) ABM61865 ck: 8069  len: 79  ! Abm61865 Propionibacterium acnes predicted
1               E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,
                  E(S)(P)(P)(P)(T)(R)(S)R(I)
       40: RPAAK  ESPPPTRSRI  SRLGR

Databases searched:
     EMBL, Release 2.0, Released on 29Jan2004, Formatted on 12Feb2004

    Total finds:          14
   Total length:  282,547,505
Total sequences:    1,586,107
       CPU time:        11:19.35
```

Number in blue corresponds to citation.

```
;  TOIG of: aab10470  check: 3609  from: 1  to: 543
;
; ID   AAB10470 standard; protein; 543 AA.
; XX
; AC   AAB10470;
; XX
; DT   11-DEC-2000  (first entry)
; XX
; DE   Shewanella putrefaciens PKS protein ORF9.
; XX
; KW   PKS pathway; polyunsaturated long chain fatty acid; plant; transgenic;
; KW   polyketide-like synthesis; PUFA; dietary supplement; intravenous feeding;
; KW   malnutrition; cooking oil; cooking fat; margarine;
; KW   docosahexenoic acid production; eicosapentenoic acid production.
; XX
; OS   Shewanella putrefaciens.
; XX
; PN   WO200042195-A2.
; XX
; PD   20-JUL-2000.
; XX
; PF   14-JAN-2000; 2000WO-US000956.
; XX
; PR   14-JAN-1999;   99US-00231899.
; XX
; PA   (CALJ ) CALGENE LLC.
; XX
; PI   Facciotti D,  Metz JG,  Lassner M;
; XX
; DR   WPI; 2000-476063/41.
; XX
; PT   New DNA sequences encoding for polyketide (PK)-like synthesis pathway
; PT   genes from Shewanella, Vibrio and Schizochtrium, useful for creating
; PT   transgenic plants that express poly-unsaturated long chain fatty acids.
; XX
; PS   Example 1; Fig 4J; 302pp; English.
; XX
; CC   This invention describes novel DNA sequences encoding for polyketide (PK)
; CC   -like synthesis (PKS-like) pathway genes from Shewanella, Vibrio and
; CC   Schizochtrium. The nucleic acids are useful for isolating related
; CC   molecules or in methods to detect organisms expressing the PKS-like
; CC   genes. They are also useful for creating transgenic plants that express
; CC   poly-unsaturated long chain fatty acids. The poly-unsaturated long chain
; CC   fatty acids produced recombinantly are useful as dietary supplements for
; CC   patients undergoing intravenous feeding or for preventing or treating
; CC   malnutrition. The poly-unsaturated long chain fatty acids can also be
; CC   incorporated into cooking oils, fats or margarine formulated so that in
; CC   normal use the recipient receives a desired amount of poly-unsaturated
; CC   long chain fatty acids. The nucleic acids are also useful in large scale
; CC   production of docosahexenoic acid and eicosapentenoic acid, and for the
; CC   modification of the fatty acid profile of host cells and edible plant
; CC   tissues and/or plant parts. Transgenic production of polyunsaturated
; CC   fatty acids in particular host cells allows quicker purification from
; CC   natural sources such as fish or plants. This sequence represents the
; CC   Shewanella putrefaciens PKS protein cluster ORF9 which is described in
; CC   the method of the invention
; XX
; SQ   Sequence 543 AA;
;
;  AAB10470  Length: 543  April 15, 2004 09:21  Type: P  Check: 3609  ..
aab10470
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: aag52216  check: 7182  from: 1  to: 274
;
; ID   AAG52216 standard; protein; 274 AA.
; XX
; AC   AAG52216;
; XX
; DT   18-OCT-2000  (first entry)
; XX
; DE   Arabidopsis thaliana protein fragment SEQ ID NO: 66352.
; XX
; KW   Protein identification; signal transduction pathway; metabolic pathway;
; KW   hybridisation assay; genetic mapping; gene expression control; promoter;
; KW   termination sequence.
; XX
; OS   Arabidopsis thaliana.
; XX
; PN   EP1033405-A2.
; XX
; PD   06-SEP-2000.
; XX
; PF   25-FEB-2000; 2000EP-00301439.
; XX
; PR   25-FEB-1999;   99US-0121825P.
; PR   05-MAR-1999;   99US-0123180P.
; PR   09-MAR-1999;   99US-0123548P.
; PR   23-MAR-1999;   99US-0125788P.
; PR   25-MAR-1999;   99US-0126264P.
; PR   29-MAR-1999;   99US-0126785P.
; PR   01-APR-1999;   99US-0127462P.
; PR   06-APR-1999;   99US-0128234P.
; PR   08-APR-1999;   99US-0128714P.
; PR   16-APR-1999;   99US-0129845P.
; PR   19-APR-1999;   99US-0130077P.
; PR   21-APR-1999;   99US-0130449P.
; PR   23-APR-1999;   99US-0130510P.
; PR   23-APR-1999;   99US-0130891P.
; PR   28-APR-1999;   99US-0131449P.
; PR   30-APR-1999;   99US-0132048P.
; PR   30-APR-1999;   99US-0132407P.
; PR   04-MAY-1999;   99US-0132484P.
; PR   05-MAY-1999;   99US-0132485P.
; PR   06-MAY-1999;   99US-0132486P.
; PR   06-MAY-1999;   99US-0132487P.
; PR   07-MAY-1999;   99US-0132863P.
; PR   11-MAY-1999;   99US-0134256P.
; PR   14-MAY-1999;   99US-0134218P.
; PR   14-MAY-1999;   99US-0134219P.
; PR   14-MAY-1999;   99US-0134221P.
; PR   14-MAY-1999;   99US-0134370P.
; PR   18-MAY-1999;   99US-0134768P.
; PR   19-MAY-1999;   99US-0134941P.
; PR   20-MAY-1999;   99US-0135124P.
; PR   21-MAY-1999;   99US-0135353P.
; PR   24-MAY-1999;   99US-0135629P.
; PR   25-MAY-1999;   99US-0136021P.
; PR   27-MAY-1999;   99US-0136392P.
; PR   28-MAY-1999;   99US-0136782P.
; PR   01-JUN-1999;   99US-0137222P.
; PR   03-JUN-1999;   99US-0137528P.
; PR   04-JUN-1999;   99US-0137502P.
; PR   07-JUN-1999;   99US-0137724P.
; PR   08-JUN-1999;   99US-0138094P.
; PR   10-JUN-1999;   99US-0138540P.
; PR   10-JUN-1999;   99US-0138847P.
; PR   14-JUN-1999;   99US-0139119P.
; PR   16-JUN-1999;   99US-0139452P.
; PR   16-JUN-1999;   99US-0139453P.
; PR   17-JUN-1999;   99US-0139492P.
; PR   18-JUN-1999;   99US-0139454P.
; PR   18-JUN-1999;   99US-0139455P.
; PR   18-JUN-1999;   99US-0139456P.
; PR   18-JUN-1999;   99US-0139457P.
```

; PR 18-JUN-1999; 99US-0139458P.
; PR 18-JUN-1999; 99US-0139459P.
; PR 18-JUN-1999; 99US-0139460P.
; PR 18-JUN-1999; 99US-0139461P.
; PR 18-JUN-1999; 99US-0139462P.
; PR 18-JUN-1999; 99US-0139463P.
; PR 18-JUN-1999; 99US-0139750P.
; PR 18-JUN-1999; 99US-0139763P.
; PR 21-JUN-1999; 99US-0139817P.
; PR 22-JUN-1999; 99US-0139899P.
; PR 23-JUN-1999; 99US-0140353P.
; PR 23-JUN-1999; 99US-0140354P.
; PR 24-JUN-1999; 99US-0140695P.
; PR 28-JUN-1999; 99US-0140823P.
; PR 29-JUN-1999; 99US-0140991P.
; PR 30-JUN-1999; 99US-0141287P.
; PR 01-JUL-1999; 99US-0141842P.
; PR 01-JUL-1999; 99US-0142154P.
; PR 02-JUL-1999; 99US-0142055P.
; PR 06-JUL-1999; 99US-0142390P.
; PR 08-JUL-1999; 99US-0142803P.
; PR 09-JUL-1999; 99US-0142920P.
; PR 12-JUL-1999; 99US-0142977P.
; PR 13-JUL-1999; 99US-0143542P.
; PR 14-JUL-1999; 99US-0143624P.
; PR 15-JUL-1999; 99US-0144005P.
; PR 16-JUL-1999; 99US-0144085P.
; PR 16-JUL-1999; 99US-0144086P.
; PR 19-JUL-1999; 99US-0144325P.
; PR 19-JUL-1999; 99US-0144331P.
; PR 19-JUL-1999; 99US-0144332P.
; PR 19-JUL-1999; 99US-0144333P.
; PR 19-JUL-1999; 99US-0144334P.
; PR 19-JUL-1999; 99US-0144335P.
; PR 20-JUL-1999; 99US-0144352P.
; PR 20-JUL-1999; 99US-0144632P.
; PR 20-JUL-1999; 99US-0144884P.
; PR 21-JUL-1999; 99US-0144814P.
; PR 21-JUL-1999; 99US-0145086P.
; PR 21-JUL-1999; 99US-0145088P.
; PR 22-JUL-1999; 99US-0145085P.
; PR 22-JUL-1999; 99US-0145087P.
; PR 22-JUL-1999; 99US-0145089P.
; PR 22-JUL-1999; 99US-0145192P.
; PR 23-JUL-1999; 99US-0145145P.
; PR 23-JUL-1999; 99US-0145218P.
; PR 23-JUL-1999; 99US-0145224P.
; PR 26-JUL-1999; 99US-0145276P.
; PR 27-JUL-1999; 99US-0145913P.
; PR 27-JUL-1999; 99US-0145918P.
; PR 27-JUL-1999; 99US-0145919P.
; PR 28-JUL-1999; 99US-0145951P.
; PR 02-AUG-1999; 99US-0146386P.
; PR 02-AUG-1999; 99US-0146388P.
; PR 02-AUG-1999; 99US-0146389P.
; PR 03-AUG-1999; 99US-0147038P.
; PR 04-AUG-1999; 99US-0147204P.
; PR 04-AUG-1999; 99US-0147302P.
; PR 05-AUG-1999; 99US-0147192P.
; PR 05-AUG-1999; 99US-0147260P.
; PR 06-AUG-1999; 99US-0147303P.
; PR 06-AUG-1999; 99US-0147416P.
; PR 09-AUG-1999; 99US-0147493P.
; PR 09-AUG-1999; 99US-0147935P.
; PR 10-AUG-1999; 99US-0148171P.
; PR 11-AUG-1999; 99US-0148319P.
; PR 12-AUG-1999; 99US-0148341P.
; PR 13-AUG-1999; 99US-0148565P.
; PR 13-AUG-1999; 99US-0148684P.
; PR 16-AUG-1999; 99US-0149368P.
; PR 17-AUG-1999; 99US-0149175P.
; PR 18-AUG-1999; 99US-0149426P.
; PR 20-AUG-1999; 99US-0149722P.
; PR 20-AUG-1999; 99US-0149723P.
; PR 20-AUG-1999; 99US-0149929P.
; PR 23-AUG-1999; 99US-0149902P.
; PR 23-AUG-1999; 99US-0149930P.
; PR 25-AUG-1999; 99US-0150566P.
; PR 26-AUG-1999; 99US-0150884P.
; PR 27-AUG-1999; 99US-0151065P.
; PR 27-AUG-1999; 99US-0151066P.
; PR 27-AUG-1999; 99US-0151080P.
; PR 30-AUG-1999; 99US-0151303P.
; PR 31-AUG-1999; 99US-0151438P.
; PR 01-SEP-1999; 99US-0151930P.
; PR 07-SEP-1999; 99US-0152363P.
; PR 10-SEP-1999; 99US-0153070P.
; PR 13-SEP-1999; 99US-0153758P.
; PR 15-SEP-1999; 99US-0154018P.
; PR 16-SEP-1999; 99US-0154039P.
; PR 20-SEP-1999; 99US-0154779P.
; PR 22-SEP-1999; 99US-0155139P.
; PR 23-SEP-1999; 99US-0155486P.
; PR 24-SEP-1999; 99US-0155659P.
; PR 28-SEP-1999; 99US-0156458P.
; PR 29-SEP-1999; 99US-0156596P.
; PR 04-OCT-1999; 99US-0157117P.
; PR 05-OCT-1999; 99US-0157753P.
; PR 06-OCT-1999; 99US-0157865P.
; PR 07-OCT-1999; 99US-0158029P.
; PR 08-OCT-1999; 99US-0158232P.
; PR 12-OCT-1999; 99US-0158369P.
; PR 13-OCT-1999; 99US-0159293P.
; PR 13-OCT-1999; 99US-0159294P.
; PR 13-OCT-1999; 99US-0159295P.
; PR 14-OCT-1999; 99US-0159329P.
; PR 14-OCT-1999; 99US-0159330P.
; PR 14-OCT-1999; 99US-0159331P.
; PR 14-OCT-1999; 99US-0159637P.
; PR 14-OCT-1999; 99US-0159638P.
; PR 18-OCT-1999; 99US-0159584P.
; PR 21-OCT-1999; 99US-0160741P.
; PR 21-OCT-1999; 99US-0160767P.
; PR 21-OCT-1999; 99US-0160768P.
; PR 21-OCT-1999; 99US-0160770P.
; PR 21-OCT-1999; 99US-0160814P.
; PR 21-OCT-1999; 99US-0160815P.
; PR 22-OCT-1999; 99US-0160980P.
; PR 22-OCT-1999; 99US-0160981P.
; PR 22-OCT-1999; 99US-0160989P.
; PR 25-OCT-1999; 99US-0161404P.
; PR 25-OCT-1999; 99US-0161405P.
; PR 25-OCT-1999; 99US-0161406P.
; PR 26-OCT-1999; 99US-0161359P.
; PR 26-OCT-1999; 99US-0161360P.
; PR 26-OCT-1999; 99US-0161361P.
; PR 28-OCT-1999; 99US-0161920P.
; PR 28-OCT-1999; 99US-0161992P.
; PR 28-OCT-1999; 99US-0161993P.
; PR 29-OCT-1999; 99US-0162142P.
; PR 29-OCT-1999; 99US-0162143P.
; PR 29-OCT-1999; 99US-0162228P.
; PR 01-NOV-1999; 99US-0162891P.
; PR 01-NOV-1999; 99US-0162894P.
; PR 01-NOV-1999; 99US-0162895P.
; PR 02-NOV-1999; 99US-0163091P.
; PR 02-NOV-1999; 99US-0163092P.
; PR 02-NOV-1999; 99US-0163093P.
; PR 03-NOV-1999; 99US-0163248P.
; PR 03-NOV-1999; 99US-0163249P.
; PR 03-NOV-1999; 99US-0163281P.
; PR 04-NOV-1999; 99US-0163379P.
; PR 04-NOV-1999; 99US-0163380P.
; PR 04-NOV-1999; 99US-0163381P.
; PR 08-NOV-1999; 99US-0164146P.
; PR 08-NOV-1999; 99US-0164150P.

; PR 08-NOV-1999; 99US-0164151P.
; PR 09-NOV-1999; 99US-0164259P.
; PR 09-NOV-1999; 99US-0164260P.
; PR 10-NOV-1999; 99US-0164317P.
; PR 10-NOV-1999; 99US-0164318P.
; PR 10-NOV-1999; 99US-0164319P.
; PR 10-NOV-1999; 99US-0164321P.
; PR 10-NOV-1999; 99US-0164544P.
; PR 10-NOV-1999; 99US-0164545P.
; PR 10-NOV-1999; 99US-0164548P.
; PR 12-NOV-1999; 99US-0164870P.
; PR 12-NOV-1999; 99US-0164871P.
; PR 12-NOV-1999; 99US-0164959P.
; PR 12-NOV-1999; 99US-0164960P.
; PR 12-NOV-1999; 99US-0164961P.
; PR 12-NOV-1999; 99US-0164962P.
; PR 15-NOV-1999; 99US-0164926P.
; PR 15-NOV-1999; 99US-0164927P.
; PR 15-NOV-1999; 99US-0164929P.
; PR 16-NOV-1999; 99US-0165661P.
; PR 16-NOV-1999; 99US-0165669P.
; PR 16-NOV-1999; 99US-0165671P.
; PR 17-NOV-1999; 99US-0165911P.
; PR 17-NOV-1999; 99US-0165918P.
; PR 17-NOV-1999; 99US-0165919P.
; PR 18-NOV-1999; 99US-0166157P.
; PR 18-NOV-1999; 99US-0166158P.
; PR 18-NOV-1999; 99US-0166173P.
; PR 19-NOV-1999; 99US-0166411P.
; PR 19-NOV-1999; 99US-0166412P.
; PR 19-NOV-1999; 99US-0166419P.
; PR 22-NOV-1999; 99US-0166733P.
; PR 22-NOV-1999; 99US-0166750P.
; PR 23-NOV-1999; 99US-0167362P.
; PR 24-NOV-1999; 99US-0167233P.
; PR 24-NOV-1999; 99US-0167234P.
; PR 24-NOV-1999; 99US-0167235P.
; PR 24-NOV-1999; 99US-0167382P.
; PR 30-NOV-1999; 99US-0167902P.
; PR 30-NOV-1999; 99US-0167904P.
; PR 30-NOV-1999; 99US-0167908P.
; PR 01-DEC-1999; 99US-0168231P.
; PR 01-DEC-1999; 99US-0168232P.
; PR 01-DEC-1999; 99US-0168233P.
; PR 02-DEC-1999; 99US-0168546P.
; PR 02-DEC-1999; 99US-0168548P.
; PR 02-DEC-1999; 99US-0168549P.
; PR 03-DEC-1999; 99US-0168673P.
; PR 03-DEC-1999; 99US-0168674P.
; PR 03-DEC-1999; 99US-0168675P.
; PR 07-DEC-1999; 99US-0169278P.
; PR 07-DEC-1999; 99US-0169298P.
; PR 07-DEC-1999; 99US-0169302P.
; PR 08-DEC-1999; 99US-0169691P.
; PR 08-DEC-1999; 99US-0169692P.
; PR 16-DEC-1999; 99US-0171098P.
; PR 16-DEC-1999; 99US-0171107P.
; PR 16-DEC-1999; 99US-0171114P.
; PR 19-JAN-2000; 2000US-0176866P.
; PR 19-JAN-2000; 2000US-0176867P.
; PR 19-JAN-2000; 2000US-0176910P.
; PR 26-JAN-2000; 2000US-0178166P.
; PR 27-JAN-2000; 2000US-0177666P.
; PR 27-JAN-2000; 2000US-0178544P.
; PR 27-JAN-2000; 2000US-0178545P.
; PR 27-JAN-2000; 2000US-0178546P.
; PR 27-JAN-2000; 2000US-0178547P.
; PR 28-JAN-2000; 2000US-0178754P.
; PR 28-JAN-2000; 2000US-0178755P.
; PR 01-FEB-2000; 2000US-0179388P.
; PR 01-FEB-2000; 2000US-0179395P.
; PR 03-FEB-2000; 2000US-0180039P.
; PR 03-FEB-2000; 2000US-0180139P.
; PR 04-FEB-2000; 2000US-0180206P.
; PR 04-FEB-2000; 2000US-0180207P.
; PR 07-FEB-2000; 2000US-0180695P.
; PR 07-FEB-2000; 2000US-0180696P.
; PR 09-FEB-2000; 2000US-0181214P.
; PR 09-FEB-2000; 2000US-0181228P.
; PR 10-FEB-2000; 2000US-0181476P.
; PR 10-FEB-2000; 2000US-0181551P.
; PR 15-FEB-2000; 2000US-0182477P.
; PR 15-FEB-2000; 2000US-0182478P.
; PR 15-FEB-2000; 2000US-0182512P.
; PR 15-FEB-2000; 2000US-0182516P.
; PR 17-FEB-2000; 2000US-0183165P.
; PR 17-FEB-2000; 2000US-0183166P.
; XX
; PA (CERE-) CERES INC.
; XX
; PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
; PI Zheng L, Dumas J;
; XX
; DR WPI; 2000-507395/46.
; DR N-PSDB; AAC50846.
; XX
; PT New sequence determined DNA fragments (SDFs) from different plant
; PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
; PT protein coding sequences, untranslated regions, or as 3' termination
; PT sequences.
; XX
; PS Claim 19; SEQ ID NO 66352; 344pp + Sequence Listing; English.
; XX
; CC The present sequence is a putative protein fragment from Arabidopsis
; CC thaliana. Its coding sequence was isolated by carrying out RT-PCR on all
; CC of the mRNA obtained from the plant, and creating a cDNA library which
; CC could then be sequenced, allowing the putative protein sequence(s) to be
; CC obtained. This sequence may be useful for protein identification and for
; CC aiding in the elucidation of signal transduction and metabolic pathways.
; CC Its coding sequence has a use in the control of gene expression as a
; CC promoter, coding sequence, 3'UTR or termination sequence, for controlling
; CC the behaviour of a gene within the chromosome, as a tool for use in
; CC genetic mapping, including a use in hybridisation assays, for recognition
; CC or isolation of similar DNA fragments, or for the identification of a
; CC particular organism
; XX
; SQ Sequence 274 AA;
;
; AAG52216 Length: 274 April 15, 2004 09:21 Type: P Check: 7182 ..
aag52216
MGDFSIQISSKLINQLAEGNDQPKRRAKKTKPKVSPQSKQKTNQDEEKKPNPVAELPMQPPFFFPIPPQG
AASTELESIKSVVKESEKVLEKLELQEKNIVREVTERAKDLREKEFKIPEPKPMPCSSDHEAWMKCYKEN
IGSPLKCSGFVKSFQDCARRSRQQVNPEENSNTLNRVNLGEQIFLSIFNVMTRMMLGAIVEEEERTILGN
ELKKLILLFQISKEAQKYSYVPHRNSKTSVVAGYTVLKDLKIFINVWAINRDTKNWEEEGSKRV1

```
;  TOIG of: aag52217  check: 376  from: 1  to: 217
;
; ID   AAG52217 standard; protein; 217 AA.
; XX
; AC   AAG52217;
; XX
; DT   18-OCT-2000  (first entry)
; XX
; DE   Arabidopsis thaliana protein fragment SEQ ID NO: 66353.
; XX
; KW   Protein identification; signal transduction pathway; metabolic pathway;
; KW   hybridisation assay; genetic mapping; gene expression control; promoter;
; KW   termination sequence.
; XX
; OS   Arabidopsis thaliana.
; XX
; PN   EP1033405-A2.
; XX
; PD   06-SEP-2000.
; XX
; PF   25-FEB-2000; 2000EP-00301439.
; XX
; PR   25-FEB-1999;   99US-0121825P.
; PR   05-MAR-1999;   99US-0123180P.
; PR   09-MAR-1999;   99US-0123548P.
; PR   23-MAR-1999;   99US-0125788P.
; PR   25-MAR-1999;   99US-0126264P.
; PR   29-MAR-1999;   99US-0126785P.
; PR   01-APR-1999;   99US-0127462P.
; PR   06-APR-1999;   99US-0128234P.
; PR   08-APR-1999;   99US-0128714P.
; PR   16-APR-1999;   99US-0129845P.
; PR   19-APR-1999;   99US-0130077P.
; PR   21-APR-1999;   99US-0130449P.
; PR   23-APR-1999;   99US-0130510P.
; PR   23-APR-1999;   99US-0130891P.
; PR   28-APR-1999;   99US-0131449P.
; PR   30-APR-1999;   99US-0132048P.
; PR   30-APR-1999;   99US-0132407P.
; PR   04-MAY-1999;   99US-0132484P.
; PR   05-MAY-1999;   99US-0132485P.
; PR   06-MAY-1999;   99US-0132486P.
; PR   06-MAY-1999;   99US-0132487P.
; PR   07-MAY-1999;   99US-0132863P.
; PR   11-MAY-1999;   99US-0134256P.
; PR   14-MAY-1999;   99US-0134218P.
; PR   14-MAY-1999;   99US-0134219P.
; PR   14-MAY-1999;   99US-0134221P.
; PR   14-MAY-1999;   99US-0134370P.
; PR   18-MAY-1999;   99US-0134768P.
; PR   19-MAY-1999;   99US-0134941P.
; PR   20-MAY-1999;   99US-0135124P.
; PR   21-MAY-1999;   99US-0135353P.
; PR   24-MAY-1999;   99US-0135629P.
; PR   25-MAY-1999;   99US-0136021P.
; PR   27-MAY-1999;   99US-0136392P.
; PR   28-MAY-1999;   99US-0136782P.
; PR   01-JUN-1999;   99US-0137222P.
; PR   03-JUN-1999;   99US-0137528P.
; PR   04-JUN-1999;   99US-0137502P.
; PR   07-JUN-1999;   99US-0137724P.
; PR   08-JUN-1999;   99US-0138094P.
; PR   10-JUN-1999;   99US-0138540P.
; PR   10-JUN-1999;   99US-0138847P.
; PR   14-JUN-1999;   99US-0139119P.
; PR   16-JUN-1999;   99US-0139452P.
; PR   16-JUN-1999;   99US-0139453P.
; PR   17-JUN-1999;   99US-0139492P.
; PR   18-JUN-1999;   99US-0139454P.
; PR   18-JUN-1999;   99US-0139455P.
; PR   18-JUN-1999;   99US-0139456P.
; PR   18-JUN-1999;   99US-0139457P.
; PR   18-JUN-1999;   99US-0139458P.
; PR   18-JUN-1999;   99US-0139459P.
; PR   18-JUN-1999;   99US-0139460P.
; PR   18-JUN-1999;   99US-0139461P.
; PR   18-JUN-1999;   99US-0139462P.
; PR   18-JUN-1999;   99US-0139463P.
; PR   18-JUN-1999;   99US-0139750P.
; PR   18-JUN-1999;   99US-0139763P.
; PR   21-JUN-1999;   99US-0139817P.
; PR   22-JUN-1999;   99US-0139899P.
; PR   23-JUN-1999;   99US-0140353P.
; PR   23-JUN-1999;   99US-0140354P.
; PR   24-JUN-1999;   99US-0140695P.
; PR   28-JUN-1999;   99US-0140823P.
; PR   29-JUN-1999;   99US-0140991P.
; PR   30-JUN-1999;   99US-0141287P.
; PR   01-JUL-1999;   99US-0141842P.
; PR   01-JUL-1999;   99US-0142154P.
; PR   02-JUL-1999;   99US-0142055P.
; PR   06-JUL-1999;   99US-0142390P.
; PR   08-JUL-1999;   99US-0142803P.
; PR   09-JUL-1999;   99US-0142920P.
; PR   12-JUL-1999;   99US-0142977P.
; PR   13-JUL-1999;   99US-0143542P.
; PR   14-JUL-1999;   99US-0143624P.
; PR   15-JUL-1999;   99US-0144005P.
; PR   16-JUL-1999;   99US-0144085P.
; PR   16-JUL-1999;   99US-0144086P.
; PR   19-JUL-1999;   99US-0144325P.
; PR   19-JUL-1999;   99US-0144331P.
; PR   19-JUL-1999;   99US-0144332P.
; PR   19-JUL-1999;   99US-0144333P.
; PR   19-JUL-1999;   99US-0144334P.
; PR   19-JUL-1999;   99US-0144335P.
; PR   20-JUL-1999;   99US-0144352P.
; PR   20-JUL-1999;   99US-0144632P.
; PR   20-JUL-1999;   99US-0144884P.
; PR   21-JUL-1999;   99US-0144814P.
; PR   21-JUL-1999;   99US-0145086P.
; PR   21-JUL-1999;   99US-0145088P.
; PR   22-JUL-1999;   99US-0145085P.
; PR   22-JUL-1999;   99US-0145087P.
; PR   22-JUL-1999;   99US-0145089P.
; PR   22-JUL-1999;   99US-0145192P.
; PR   23-JUL-1999;   99US-0145145P.
; PR   23-JUL-1999;   99US-0145218P.
; PR   23-JUL-1999;   99US-0145224P.
; PR   26-JUL-1999;   99US-0145276P.
; PR   27-JUL-1999;   99US-0145913P.
; PR   27-JUL-1999;   99US-0145918P.
; PR   27-JUL-1999;   99US-0145919P.
; PR   28-JUL-1999;   99US-0145951P.
; PR   02-AUG-1999;   99US-0146386P.
; PR   02-AUG-1999;   99US-0146388P.
; PR   02-AUG-1999;   99US-0146389P.
; PR   03-AUG-1999;   99US-0147038P.
; PR   04-AUG-1999;   99US-0147204P.
; PR   04-AUG-1999;   99US-0147302P.
; PR   05-AUG-1999;   99US-0147192P.
; PR   05-AUG-1999;   99US-0147260P.
; PR   06-AUG-1999;   99US-0147303P.
; PR   06-AUG-1999;   99US-0147416P.
; PR   09-AUG-1999;   99US-0147493P.
; PR   09-AUG-1999;   99US-0147935P.
; PR   10-AUG-1999;   99US-0148171P.
; PR   11-AUG-1999;   99US-0148319P.
; PR   12-AUG-1999;   99US-0148341P.
; PR   13-AUG-1999;   99US-0148565P.
; PR   13-AUG-1999;   99US-0148684P.
; PR   16-AUG-1999;   99US-0149368P.
; PR   17-AUG-1999;   99US-0149175P.
; PR   18-AUG-1999;   99US-0149426P.
; PR   20-AUG-1999;   99US-0149722P.
```

; PR 20-AUG-1999; 99US-0149723P.
; PR 20-AUG-1999; 99US-0149929P.
; PR 23-AUG-1999; 99US-0149902P.
; PR 23-AUG-1999; 99US-0149930P.
; PR 25-AUG-1999; 99US-0150566P.
; PR 26-AUG-1999; 99US-0150884P.
; PR 27-AUG-1999; 99US-0151065P.
; PR 27-AUG-1999; 99US-0151066P.
; PR 27-AUG-1999; 99US-0151080P.
; PR 30-AUG-1999; 99US-0151303P.
; PR 31-AUG-1999; 99US-0151438P.
; PR 01-SEP-1999; 99US-0151930P.
; PR 07-SEP-1999; 99US-0152363P.
; PR 10-SEP-1999; 99US-0153070P.
; PR 13-SEP-1999; 99US-0153758P.
; PR 15-SEP-1999; 99US-0154018P.
; PR 16-SEP-1999; 99US-0154039P.
; PR 20-SEP-1999; 99US-0154779P.
; PR 22-SEP-1999; 99US-0155139P.
; PR 23-SEP-1999; 99US-0155486P.
; PR 24-SEP-1999; 99US-0155659P.
; PR 28-SEP-1999; 99US-0156458P.
; PR 29-SEP-1999; 99US-0156596P.
; PR 04-OCT-1999; 99US-0157117P.
; PR 05-OCT-1999; 99US-0157753P.
; PR 06-OCT-1999; 99US-0157865P.
; PR 07-OCT-1999; 99US-0158029P.
; PR 08-OCT-1999; 99US-0158232P.
; PR 12-OCT-1999; 99US-0158369P.
; PR 13-OCT-1999; 99US-0159293P.
; PR 13-OCT-1999; 99US-0159294P.
; PR 13-OCT-1999; 99US-0159295P.
; PR 14-OCT-1999; 99US-0159329P.
; PR 14-OCT-1999; 99US-0159330P.
; PR 14-OCT-1999; 99US-0159331P.
; PR 14-OCT-1999; 99US-0159637P.
; PR 14-OCT-1999; 99US-0159638P.
; PR 18-OCT-1999; 99US-0159584P.
; PR 21-OCT-1999; 99US-0160741P.
; PR 21-OCT-1999; 99US-0160767P.
; PR 21-OCT-1999; 99US-0160768P.
; PR 21-OCT-1999; 99US-0160770P.
; PR 21-OCT-1999; 99US-0160814P.
; PR 21-OCT-1999; 99US-0160815P.
; PR 22-OCT-1999; 99US-0160980P.
; PR 22-OCT-1999; 99US-0160981P.
; PR 22-OCT-1999; 99US-0160989P.
; PR 25-OCT-1999; 99US-0161404P.
; PR 25-OCT-1999; 99US-0161405P.
; PR 25-OCT-1999; 99US-0161406P.
; PR 26-OCT-1999; 99US-0161359P.
; PR 26-OCT-1999; 99US-0161360P.
; PR 26-OCT-1999; 99US-0161361P.
; PR 28-OCT-1999; 99US-0161920P.
; PR 28-OCT-1999; 99US-0161992P.
; PR 28-OCT-1999; 99US-0161993P.
; PR 29-OCT-1999; 99US-0162142P.
; PR 29-OCT-1999; 99US-0162143P.
; PR 29-OCT-1999; 99US-0162228P.
; PR 01-NOV-1999; 99US-0162891P.
; PR 01-NOV-1999; 99US-0162894P.
; PR 01-NOV-1999; 99US-0162895P.
; PR 02-NOV-1999; 99US-0163091P.
; PR 02-NOV-1999; 99US-0163092P.
; PR 02-NOV-1999; 99US-0163093P.
; PR 03-NOV-1999; 99US-0163248P.
; PR 03-NOV-1999; 99US-0163249P.
; PR 03-NOV-1999; 99US-0163281P.
; PR 04-NOV-1999; 99US-0163379P.
; PR 04-NOV-1999; 99US-0163380P.
; PR 04-NOV-1999; 99US-0163381P.
; PR 08-NOV-1999; 99US-0164146P.
; PR 08-NOV-1999; 99US-0164150P.
; PR 08-NOV-1999; 99US-0164151P.
; PR 09-NOV-1999; 99US-0164259P.
; PR 09-NOV-1999; 99US-0164260P.
; PR 10-NOV-1999; 99US-0164317P.
; PR 10-NOV-1999; 99US-0164318P.
; PR 10-NOV-1999; 99US-0164319P.
; PR 10-NOV-1999; 99US-0164321P.
; PR 10-NOV-1999; 99US-0164544P.
; PR 10-NOV-1999; 99US-0164545P.
; PR 10-NOV-1999; 99US-0164548P.
; PR 12-NOV-1999; 99US-0164870P.
; PR 12-NOV-1999; 99US-0164871P.
; PR 12-NOV-1999; 99US-0164959P.
; PR 12-NOV-1999; 99US-0164960P.
; PR 12-NOV-1999; 99US-0164961P.
; PR 12-NOV-1999; 99US-0164962P.
; PR 15-NOV-1999; 99US-0164926P.
; PR 15-NOV-1999; 99US-0164927P.
; PR 15-NOV-1999; 99US-0164929P.
; PR 16-NOV-1999; 99US-0165661P.
; PR 16-NOV-1999; 99US-0165669P.
; PR 16-NOV-1999; 99US-0165671P.
; PR 17-NOV-1999; 99US-0165911P.
; PR 17-NOV-1999; 99US-0165918P.
; PR 17-NOV-1999; 99US-0165919P.
; PR 18-NOV-1999; 99US-0166157P.
; PR 18-NOV-1999; 99US-0166158P.
; PR 18-NOV-1999; 99US-0166173P.
; PR 19-NOV-1999; 99US-0166411P.
; PR 19-NOV-1999; 99US-0166412P.
; PR 19-NOV-1999; 99US-0166419P.
; PR 22-NOV-1999; 99US-0166733P.
; PR 22-NOV-1999; 99US-0166750P.
; PR 23-NOV-1999; 99US-0167362P.
; PR 24-NOV-1999; 99US-0167233P.
; PR 24-NOV-1999; 99US-0167234P.
; PR 24-NOV-1999; 99US-0167235P.
; PR 24-NOV-1999; 99US-0167382P.
; PR 30-NOV-1999; 99US-0167902P.
; PR 30-NOV-1999; 99US-0167904P.
; PR 30-NOV-1999; 99US-0167908P.
; PR 01-DEC-1999; 99US-0168231P.
; PR 01-DEC-1999; 99US-0168232P.
; PR 01-DEC-1999; 99US-0168233P.
; PR 02-DEC-1999; 99US-0168546P.
; PR 02-DEC-1999; 99US-0168548P.
; PR 02-DEC-1999; 99US-0168549P.
; PR 03-DEC-1999; 99US-0168673P.
; PR 03-DEC-1999; 99US-0168674P.
; PR 03-DEC-1999; 99US-0168675P.
; PR 07-DEC-1999; 99US-0169278P.
; PR 07-DEC-1999; 99US-0169298P.
; PR 07-DEC-1999; 99US-0169302P.
; PR 08-DEC-1999; 99US-0169691P.
; PR 08-DEC-1999; 99US-0169692P.
; PR 16-DEC-1999; 99US-0171098P.
; PR 16-DEC-1999; 99US-0171107P.
; PR 16-DEC-1999; 99US-0171114P.
; PR 19-JAN-2000; 2000US-0176866P.
; PR 19-JAN-2000; 2000US-0176867P.
; PR 19-JAN-2000; 2000US-0176910P.
; PR 26-JAN-2000; 2000US-0178166P.
; PR 27-JAN-2000; 2000US-0177666P.
; PR 27-JAN-2000; 2000US-0178544P.
; PR 27-JAN-2000; 2000US-0178545P.
; PR 27-JAN-2000; 2000US-0178546P.
; PR 27-JAN-2000; 2000US-0178547P.
; PR 28-JAN-2000; 2000US-0178754P.
; PR 28-JAN-2000; 2000US-0178755P.
; PR 01-FEB-2000; 2000US-0179388P.
; PR 01-FEB-2000; 2000US-0179395P.
; PR 03-FEB-2000; 2000US-0180039P.
; PR 03-FEB-2000; 2000US-0180139P.

```
; PR   04-FEB-2000; 2000US-0180206P.
; PR   04-FEB-2000; 2000US-0180207P.
; PR   07-FEB-2000; 2000US-0180695P.
; PR   07-FEB-2000; 2000US-0180696P.
; PR   09-FEB-2000; 2000US-0181214P.
; PR   09-FEB-2000; 2000US-0181228P.
; PR   10-FEB-2000; 2000US-0181476P.
; PR   10-FEB-2000; 2000US-0181551P.
; PR   15-FEB-2000; 2000US-0182477P.
; PR   15-FEB-2000; 2000US-0182478P.
; PR   15-FEB-2000; 2000US-0182512P.
; PR   15-FEB-2000; 2000US-0182516P.
; PR   17-FEB-2000; 2000US-0183165P.
; PR   17-FEB-2000; 2000US-0183166P.
; XX
; PA   (CERE-) CERES INC.
; XX
; PI   Alexandrov N,  Brover V,  Chen X,  Subramanian G,  Troukhan ME;
; PI   Zheng L,  Dumas J;
; XX
; DR   WPI; 2000-507395/46.
; DR   N-PSDB; AAC50846.
; XX
; PT   New sequence determined DNA fragments (SDFs) from different plant
; PT   species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
; PT   protein coding sequences, untranslated regions, or as 3' termination
; PT   sequences.
; XX
; PS   Claim 19; SEQ ID NO 66353; 344pp + Sequence Listing; English.
; XX
; CC   The present sequence is a putative protein fragment from Arabidopsis
; CC   thaliana. Its coding sequence was isolated by carrying out RT-PCR on all
; CC   of the mRNA obtained from the plant, and creating a cDNA library which
; CC   could then be sequenced, allowing the putative protein sequence(s) to be
; CC   obtained. This sequence may be useful for protein identification and for
; CC   aiding in the elucidation of signal transduction and metabolic pathways.
; CC   Its coding sequence has a use in the control of gene expression as a
; CC   promoter, coding sequence, 3'UTR or termination sequence, for controlling
; CC   the behaviour of a gene within the chromosome, as a tool for use in
; CC   genetic mapping, including a use in hybridisation assays, for recognition
; CC   or isolation of similar DNA fragments, or for the identification of a
; CC   particular organism
; XX
; SQ   Sequence 217 AA;
;
;   AAG52217  Length: 217  April 15, 2004 09:22  Type: P  Check: 376  ..
aag52217
MQPPFFPPIPPQGAASTELESIKSVVKESEKVLEKLELQEKNIVREVTERAKDLREKEFKIPEPKPMPCS
SDHEAWMKCYKENIGSPLKCSGFVKSFQDCARRSRQQVNPEENSNTLNRVNLGEQIFLSIFNVMTRMMLG
AIVEEEERTILGNELKKLILLFQISKEAQKYSYVPHRNSKTSVVAGYTVLKDLKIFINVWAINRDTKNWE
EEGSKRV1
```

```
;  TOIG of: aag52218  check: 9159  from: 1  to: 151
;
; ID   AAG52218 standard; protein; 151 AA.
; XX
; AC   AAG52218;
; XX
; DT   18-OCT-2000  (first entry)
; XX
; DE   Arabidopsis thaliana protein fragment SEQ ID NO: 66354.
; XX
; KW   Protein identification; signal transduction pathway; metabolic pathway;
; KW   hybridisation assay; genetic mapping; gene expression control; promoter;
; KW   termination sequence.
; XX
; OS   Arabidopsis thaliana.
; XX
; PN   EP1033405-A2.
; XX
; PD   06-SEP-2000.
; XX
; PF   25-FEB-2000; 2000EP-00301439.
; XX
; PR   25-FEB-1999;   99US-0121825P.
; PR   05-MAR-1999;   99US-0123180P.
; PR   09-MAR-1999;   99US-0123548P.
; PR   23-MAR-1999;   99US-0125788P.
; PR   25-MAR-1999;   99US-0126264P.
; PR   29-MAR-1999;   99US-0126785P.
; PR   01-APR-1999;   99US-0127462P.
; PR   06-APR-1999;   99US-0128234P.
; PR   08-APR-1999;   99US-0128714P.
; PR   16-APR-1999;   99US-0129845P.
; PR   19-APR-1999;   99US-0130077P.
; PR   21-APR-1999;   99US-0130449P.
; PR   23-APR-1999;   99US-0130510P.
; PR   23-APR-1999;   99US-0130891P.
; PR   28-APR-1999;   99US-0131449P.
; PR   30-APR-1999;   99US-0132048P.
; PR   30-APR-1999;   99US-0132407P.
; PR   04-MAY-1999;   99US-0132484P.
; PR   05-MAY-1999;   99US-0132485P.
; PR   06-MAY-1999;   99US-0132486P.
; PR   06-MAY-1999;   99US-0132487P.
; PR   07-MAY-1999;   99US-0132863P.
; PR   11-MAY-1999;   99US-0134256P.
; PR   14-MAY-1999;   99US-0134218P.
; PR   14-MAY-1999;   99US-0134219P.
; PR   14-MAY-1999;   99US-0134221P.
; PR   14-MAY-1999;   99US-0134370P.
; PR   18-MAY-1999;   99US-0134768P.
; PR   19-MAY-1999;   99US-0134941P.
; PR   20-MAY-1999;   99US-0135124P.
; PR   21-MAY-1999;   99US-0135353P.
; PR   24-MAY-1999;   99US-0135629P.
; PR   25-MAY-1999;   99US-0136021P.
; PR   27-MAY-1999;   99US-0136392P.
; PR   28-MAY-1999;   99US-0136782P.
; PR   01-JUN-1999;   99US-0137222P.
; PR   03-JUN-1999;   99US-0137528P.
; PR   04-JUN-1999;   99US-0137502P.
; PR   07-JUN-1999;   99US-0137724P.
; PR   08-JUN-1999;   99US-0138094P.
; PR   10-JUN-1999;   99US-0138540P.
; PR   10-JUN-1999;   99US-0138847P.
; PR   14-JUN-1999;   99US-0139119P.
; PR   16-JUN-1999;   99US-0139452P.
; PR   16-JUN-1999;   99US-0139453P.
; PR   17-JUN-1999;   99US-0139492P.
; PR   18-JUN-1999;   99US-0139454P.
; PR   18-JUN-1999;   99US-0139455P.
; PR   18-JUN-1999;   99US-0139456P.
; PR   18-JUN-1999;   99US-0139457P.
```

; PR 18-JUN-1999; 99US-0139458P.
; PR 18-JUN-1999; 99US-0139459P.
; PR 18-JUN-1999; 99US-0139460P.
; PR 18-JUN-1999; 99US-0139461P.
; PR 18-JUN-1999; 99US-0139462P.
; PR 18-JUN-1999; 99US-0139463P.
; PR 18-JUN-1999; 99US-0139750P.
; PR 18-JUN-1999; 99US-0139763P.
; PR 21-JUN-1999; 99US-0139817P.
; PR 22-JUN-1999; 99US-0139899P.
; PR 23-JUN-1999; 99US-0140353P.
; PR 23-JUN-1999; 99US-0140354P.
; PR 24-JUN-1999; 99US-0140695P.
; PR 28-JUN-1999; 99US-0140823P.
; PR 29-JUN-1999; 99US-0140991P.
; PR 30-JUN-1999; 99US-0141287P.
; PR 01-JUL-1999; 99US-0141842P.
; PR 01-JUL-1999; 99US-0142154P.
; PR 02-JUL-1999; 99US-0142055P.
; PR 06-JUL-1999; 99US-0142390P.
; PR 08-JUL-1999; 99US-0142803P.
; PR 09-JUL-1999; 99US-0142920P.
; PR 12-JUL-1999; 99US-0142977P.
; PR 13-JUL-1999; 99US-0143542P.
; PR 14-JUL-1999; 99US-0143624P.
; PR 15-JUL-1999; 99US-0144005P.
; PR 16-JUL-1999; 99US-0144085P.
; PR 16-JUL-1999; 99US-0144086P.
; PR 19-JUL-1999; 99US-0144325P.
; PR 19-JUL-1999; 99US-0144331P.
; PR 19-JUL-1999; 99US-0144332P.
; PR 19-JUL-1999; 99US-0144333P.
; PR 19-JUL-1999; 99US-0144334P.
; PR 19-JUL-1999; 99US-0144335P.
; PR 20-JUL-1999; 99US-0144352P.
; PR 20-JUL-1999; 99US-0144632P.
; PR 20-JUL-1999; 99US-0144884P.
; PR 21-JUL-1999; 99US-0144814P.
; PR 21-JUL-1999; 99US-0145086P.
; PR 21-JUL-1999; 99US-0145088P.
; PR 22-JUL-1999; 99US-0145085P.
; PR 22-JUL-1999; 99US-0145087P.
; PR 22-JUL-1999; 99US-0145089P.
; PR 22-JUL-1999; 99US-0145192P.
; PR 23-JUL-1999; 99US-0145145P.
; PR 23-JUL-1999; 99US-0145218P.
; PR 23-JUL-1999; 99US-0145224P.
; PR 26-JUL-1999; 99US-0145276P.
; PR 27-JUL-1999; 99US-0145913P.
; PR 27-JUL-1999; 99US-0145918P.
; PR 27-JUL-1999; 99US-0145919P.
; PR 28-JUL-1999; 99US-0145951P.
; PR 02-AUG-1999; 99US-0146386P.
; PR 02-AUG-1999; 99US-0146388P.
; PR 02-AUG-1999; 99US-0146389P.
; PR 03-AUG-1999; 99US-0147038P.
; PR 04-AUG-1999; 99US-0147204P.
; PR 04-AUG-1999; 99US-0147302P.
; PR 05-AUG-1999; 99US-0147192P.
; PR 05-AUG-1999; 99US-0147260P.
; PR 06-AUG-1999; 99US-0147303P.
; PR 06-AUG-1999; 99US-0147416P.
; PR 09-AUG-1999; 99US-0147493P.
; PR 09-AUG-1999; 99US-0147935P.
; PR 10-AUG-1999; 99US-0148171P.
; PR 11-AUG-1999; 99US-0148319P.
; PR 12-AUG-1999; 99US-0148341P.
; PR 13-AUG-1999; 99US-0148565P.
; PR 13-AUG-1999; 99US-0148684P.
; PR 16-AUG-1999; 99US-0149368P.
; PR 17-AUG-1999; 99US-0149175P.
; PR 18-AUG-1999; 99US-0149426P.
; PR 20-AUG-1999; 99US-0149722P.
; PR 20-AUG-1999; 99US-0149723P.
; PR 20-AUG-1999; 99US-0149929P.
; PR 23-AUG-1999; 99US-0149902P.
; PR 23-AUG-1999; 99US-0149930P.
; PR 25-AUG-1999; 99US-0150566P.
; PR 26-AUG-1999; 99US-0150884P.
; PR 27-AUG-1999; 99US-0151065P.
; PR 27-AUG-1999; 99US-0151066P.
; PR 27-AUG-1999; 99US-0151080P.
; PR 30-AUG-1999; 99US-0151303P.
; PR 31-AUG-1999; 99US-0151438P.
; PR 01-SEP-1999; 99US-0151930P.
; PR 07-SEP-1999; 99US-0152363P.
; PR 10-SEP-1999; 99US-0153070P.
; PR 13-SEP-1999; 99US-0153758P.
; PR 15-SEP-1999; 99US-0154018P.
; PR 16-SEP-1999; 99US-0154039P.
; PR 20-SEP-1999; 99US-0154779P.
; PR 22-SEP-1999; 99US-0155139P.
; PR 23-SEP-1999; 99US-0155486P.
; PR 24-SEP-1999; 99US-0155659P.
; PR 28-SEP-1999; 99US-0156458P.
; PR 29-SEP-1999; 99US-0156596P.
; PR 04-OCT-1999; 99US-0157117P.
; PR 05-OCT-1999; 99US-0157753P.
; PR 06-OCT-1999; 99US-0157865P.
; PR 07-OCT-1999; 99US-0158029P.
; PR 08-OCT-1999; 99US-0158232P.
; PR 12-OCT-1999; 99US-0158369P.
; PR 13-OCT-1999; 99US-0159293P.
; PR 13-OCT-1999; 99US-0159294P.
; PR 13-OCT-1999; 99US-0159295P.
; PR 14-OCT-1999; 99US-0159329P.
; PR 14-OCT-1999; 99US-0159330P.
; PR 14-OCT-1999; 99US-0159331P.
; PR 14-OCT-1999; 99US-0159637P.
; PR 14-OCT-1999; 99US-0159638P.
; PR 18-OCT-1999; 99US-0159584P.
; PR 21-OCT-1999; 99US-0160741P.
; PR 21-OCT-1999; 99US-0160767P.
; PR 21-OCT-1999; 99US-0160768P.
; PR 21-OCT-1999; 99US-0160770P.
; PR 21-OCT-1999; 99US-0160814P.
; PR 21-OCT-1999; 99US-0160815P.
; PR 22-OCT-1999; 99US-0160980P.
; PR 22-OCT-1999; 99US-0160981P.
; PR 22-OCT-1999; 99US-0160989P.
; PR 25-OCT-1999; 99US-0161404P.
; PR 25-OCT-1999; 99US-0161405P.
; PR 25-OCT-1999; 99US-0161406P.
; PR 26-OCT-1999; 99US-0161359P.
; PR 26-OCT-1999; 99US-0161360P.
; PR 26-OCT-1999; 99US-0161361P.
; PR 28-OCT-1999; 99US-0161920P.
; PR 28-OCT-1999; 99US-0161992P.
; PR 28-OCT-1999; 99US-0161993P.
; PR 29-OCT-1999; 99US-0162142P.
; PR 29-OCT-1999; 99US-0162143P.
; PR 29-OCT-1999; 99US-0162228P.
; PR 01-NOV-1999; 99US-0162891P.
; PR 01-NOV-1999; 99US-0162894P.
; PR 01-NOV-1999; 99US-0162895P.
; PR 02-NOV-1999; 99US-0163091P.
; PR 02-NOV-1999; 99US-0163092P.
; PR 02-NOV-1999; 99US-0163093P.
; PR 03-NOV-1999; 99US-0163248P.
; PR 03-NOV-1999; 99US-0163249P.
; PR 03-NOV-1999; 99US-0163281P.
; PR 04-NOV-1999; 99US-0163379P.
; PR 04-NOV-1999; 99US-0163380P.
; PR 04-NOV-1999; 99US-0163381P.
; PR 08-NOV-1999; 99US-0164146P.
; PR 08-NOV-1999; 99US-0164150P.

; PR 08-NOV-1999; 99US-0164151P.
; PR 09-NOV-1999; 99US-0164259P.
; PR 09-NOV-1999; 99US-0164260P.
; PR 10-NOV-1999; 99US-0164317P.
; PR 10-NOV-1999; 99US-0164318P.
; PR 10-NOV-1999; 99US-0164319P.
; PR 10-NOV-1999; 99US-0164321P.
; PR 10-NOV-1999; 99US-0164544P.
; PR 10-NOV-1999; 99US-0164545P.
; PR 10-NOV-1999; 99US-0164548P.
; PR 12-NOV-1999; 99US-0164870P.
; PR 12-NOV-1999; 99US-0164871P.
; PR 12-NOV-1999; 99US-0164959P.
; PR 12-NOV-1999; 99US-0164960P.
; PR 12-NOV-1999; 99US-0164961P.
; PR 12-NOV-1999; 99US-0164962P.
; PR 15-NOV-1999; 99US-0164926P.
; PR 15-NOV-1999; 99US-0164927P.
; PR 15-NOV-1999; 99US-0164929P.
; PR 16-NOV-1999; 99US-0165661P.
; PR 16-NOV-1999; 99US-0165669P.
; PR 16-NOV-1999; 99US-0165671P.
; PR 17-NOV-1999; 99US-0165911P.
; PR 17-NOV-1999; 99US-0165918P.
; PR 17-NOV-1999; 99US-0165919P.
; PR 18-NOV-1999; 99US-0166157P.
; PR 18-NOV-1999; 99US-0166158P.
; PR 18-NOV-1999; 99US-0166173P.
; PR 19-NOV-1999; 99US-0166411P.
; PR 19-NOV-1999; 99US-0166412P.
; PR 19-NOV-1999; 99US-0166419P.
; PR 22-NOV-1999; 99US-0166733P.
; PR 22-NOV-1999; 99US-0166750P.
; PR 23-NOV-1999; 99US-0167362P.
; PR 24-NOV-1999; 99US-0167233P.
; PR 24-NOV-1999; 99US-0167234P.
; PR 24-NOV-1999; 99US-0167235P.
; PR 24-NOV-1999; 99US-0167382P.
; PR 30-NOV-1999; 99US-0167902P.
; PR 30-NOV-1999; 99US-0167904P.
; PR 30-NOV-1999; 99US-0167908P.
; PR 01-DEC-1999; 99US-0168231P.
; PR 01-DEC-1999; 99US-0168232P.
; PR 01-DEC-1999; 99US-0168233P.
; PR 02-DEC-1999; 99US-0168546P.
; PR 02-DEC-1999; 99US-0168548P.
; PR 02-DEC-1999; 99US-0168549P.
; PR 03-DEC-1999; 99US-0168673P.
; PR 03-DEC-1999; 99US-0168674P.
; PR 03-DEC-1999; 99US-0168675P.
; PR 07-DEC-1999; 99US-0169278P.
; PR 07-DEC-1999; 99US-0169298P.
; PR 07-DEC-1999; 99US-0169302P.
; PR 08-DEC-1999; 99US-0169691P.
; PR 08-DEC-1999; 99US-0169692P.
; PR 16-DEC-1999; 99US-0171098P.
; PR 16-DEC-1999; 99US-0171107P.
; PR 16-DEC-1999; 99US-0171114P.
; PR 19-JAN-2000; 2000US-0176866P.
; PR 19-JAN-2000; 2000US-0176867P.
; PR 19-JAN-2000; 2000US-0176910P.
; PR 26-JAN-2000; 2000US-0178166P.
; PR 27-JAN-2000; 2000US-0177666P.
; PR 27-JAN-2000; 2000US-0178544P.
; PR 27-JAN-2000; 2000US-0178545P.
; PR 27-JAN-2000; 2000US-0178546P.
; PR 27-JAN-2000; 2000US-0178547P.
; PR 28-JAN-2000; 2000US-0178754P.
; PR 28-JAN-2000; 2000US-0178755P.
; PR 01-FEB-2000; 2000US-0179388P.
; PR 01-FEB-2000; 2000US-0179395P.
; PR 03-FEB-2000; 2000US-0180039P.
; PR 03-FEB-2000; 2000US-0180139P.
; PR 04-FEB-2000; 2000US-0180206P.
; PR 04-FEB-2000; 2000US-0180207P.
; PR 07-FEB-2000; 2000US-0180695P.
; PR 07-FEB-2000; 2000US-0180696P.
; PR 09-FEB-2000; 2000US-0181214P.
; PR 09-FEB-2000; 2000US-0181228P.
; PR 10-FEB-2000; 2000US-0181476P.
; PR 10-FEB-2000; 2000US-0181551P.
; PR 15-FEB-2000; 2000US-0182477P.
; PR 15-FEB-2000; 2000US-0182478P.
; PR 15-FEB-2000; 2000US-0182512P.
; PR 15-FEB-2000; 2000US-0182516P.
; PR 17-FEB-2000; 2000US-0183165P.
; PR 17-FEB-2000; 2000US-0183166P.
; XX
; PA (CERE-) CERES INC.
; XX
; PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
; PI Zheng L, Dumas J;
; XX
; DR WPI; 2000-507395/46.
; DR N-PSDB; AAC50846.
; XX
; PT New sequence determined DNA fragments (SDFs) from different plant
; PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
; PT protein coding sequences, untranslated regions, or as 3' termination
; PT sequences.
; XX
; PS Claim 19; SEQ ID NO 66354; 344pp + Sequence Listing; English.
; XX
; CC The present sequence is a putative protein fragment from Arabidopsis
; CC thaliana. Its coding sequence was isolated by carrying out RT-PCR on all
; CC of the mRNA obtained from the plant, and creating a cDNA library which
; CC could then be sequenced, allowing the putative protein sequence(s) to be
; CC obtained. This sequence may be useful for protein identification and for
; CC aiding in the elucidation of signal transduction and metabolic pathways.
; CC Its coding sequence has a use in the control of gene expression as a
; CC promoter, coding sequence, 3'UTR or termination sequence, for controlling
; CC the behaviour of a gene within the chromosome, as a tool for use in
; CC genetic mapping, including a use in hybridisation assays, for recognition
; CC or isolation of similar DNA fragments, or for the identification of a
; CC particular organism
; XX
; SQ Sequence 151 AA;
;
; AAG52218 Length: 151 April 15, 2004 09:22 Type: P Check: 9159 ..
aag52218
MPCSSDHEAWMKCYKENIGSPLKCSGFVKSFQDCARRSRQQVNPEENSNTLNRVNLGEQIFLSIFNVMTR
MMLGAIVEEEERTILGNELKKLILLFQISKEAQKYSYVPHRNSKTSVVAGYTVLKDLKIFINVWAINRDT
KNWEEEGSKRV1

```
;  TOIG of: aag64458  check: 9499  from: 1  to: 542
;
; ID   AAG64458 standard; protein; 542 AA.
; XX
; AC   AAG64458;
; XX
; DT   22-OCT-2001  (first entry)
; XX
; DE   S. putrefaciens eicosapentanoic acid synthesis enzyme 6.
; XX
; KW   Cyanobacterium; eicosapentanoic acid; EPA; plasmid.
; XX
; OS   Shewanella putrefaciens.
; XX
; PN   JP2001145490-A.
; XX
; PD   29-MAY-2001.
; XX
; PF   19-NOV-1999;   99JP-00329169.
; XX
; PR   19-NOV-1999;   99JP-00329169.
; XX
; PA   (SAGA ) SAGAMI CHEM RES CENT.
; PA   (BIOI-) BIOINDUSTRY KYOKAI SH.
; PA   (KEIZ-) KEIZAI SANGYOSHO SANGYO GIJUTSU SOGO KEN.
; XX
; DR   WPI; 2001-406151/43.
; DR   N-PSDB; AAH47839.
; XX
; PT   A plasmid in which eicosapentaenoic acid biosynthesis gene group is
; PT   cloned and used to transform cyanobacterium so that it produces
; PT   eicosapentaenoic acid.
; XX
; PS   Claim 2; Page 60-61; 62pp; Japanese.
; XX
; CC   The invention relates to a plasmid prepared by cloning the Shewanella
; CC   putrefaciens SCRC-2874 (FERM BP-1625) eicosapentanoic acid (EPA)
; CC   synthesis gene cluster (AAH47833) into a broad host vector. The plasmid
; CC   is used to transform cyanobacterium and produce EPA. The present sequence
; CC   is that of a EPA biosynthesis enzyme of the invention
; XX
; SQ   Sequence 542 AA;
;
;  AAG64458  Length: 542  April 15, 2004 09:22  Type: P  Check: 9499  ..
aag64458
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCSFGAAGLIPSRVEAAIN
RIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVVV
GNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVTL
LPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLAT
TEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDEI
WAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAKG
SYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: aar30729  check: 500  from: 1  to: 870
;
; ID   AAR30729 standard; protein; 870 AA.
; XX
; AC   AAR30729;
; XX
; DT   27-AUG-2003  (revised)
; DT   25-MAR-2003  (revised)
; DT   20-MAY-1993  (first entry)
; XX
; DE   p100 protein from human herpes virus type 6.
; XX
; KW   antibodys; monoclonal antibody; ELISA assays; CMV; cytomegalovirus.
; XX
; OS   Human herpesvirus 6.
; XX
; PN   EP524421-A1.
; XX
; PD   27-JAN-1993.
; XX
; PF   15-JUN-1992;   92EP-00110047.
; XX
; PR   08-JUL-1991;   91EP-00111338.
; XX
; PA   (BEHW ) BEHRINGWERKE AG.
; XX
; PI   Neipel F,  Fleckenstein B;
; XX
; DR   WPI; 1993-028531/04.
; DR   P-PSDB; AAR30729.
; XX
; PT   Human herpes virus type 6 protein p100 DNA sequence - useful in
; PT   prophylaxis, treatment and differential diagnosis of human herpes virus-6
; PT   infections.
; XX
; PS   Claim 1; Page 12; 25pp; English.
; XX
; CC   This sequence is the p100 protein from human herpes virus type6. The
; CC   protein and antibodies to it can be used for treatment or prevention of
; CC   HHV-6 infections. The DNA , protein and Ab are also useful in eg. ELISA
; CC   assays esp. for differentiating between HHV-6 and cytomegalovirus
; CC   infections. These assays are more sensitive and specific than
; CC   immunofluorescence methods currently used. (Updated on 25-MAR-2003 to
; CC   correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
; XX
; SQ   Sequence 870 AA;
;
;  AAR30729  Length: 870  April 15, 2004 09:20  Type: P  Check: 500  ..
aar30729
MDLQRHPIPFAWLDRDKVERLTDFLSNLERLDNVDLREHPHVTNSCVVREGDDVDDLKTLYNLLVLWLMY
HYVLSKRKPDYNAIWQDITKLQSVVNEYLNSKGLNKGIFENMFTNKEKFESQFSDINRALLRLGNFIKWG
SNVAIDTPYVNLTAEDSSEIENNLQDAEKNMLWYTVYNINDPWDENGYLITSINKLTYLGKLFLALTQSW
SKLEKVAMSQIVITQNHLSGHLRRHDNFNIVYSHRVLQTPLTGQRVESFLKIITSDYDIIKSSLESHSAS
KAFSMSEIGPNSLMDFVPLRGDIHSNLTLPSMSIDTKKSSLDPARLKKSNSRSLDSFLRMQRQPKFLELD
SVDNAGEKILLKEATLGGENVKATTPASSVSLMSGVESPSSFTSTNLDLPLSSFTSTNLDLRDKSHGNYK
IGPSGILDFNVKFPPNAQLNTNGVDLLQDKTSIGSPSSGITDVVNGFANLNLHQNKSNVSPPWSRNTAAN
ADFLDPVHRFVPEQTGTPFVLNNSDVAGSEAKHTTYSTETGVSPRNVFLIKDLRGKDGFRKQKQSDIPKS
LTKERNDKAIMHSREVTGDSGDATETVGARNSPALRKIKQANDFFAGLNKKNDRDVLRGGKGNSKDLHSG
GNAKKKEMSGKFNDDKEMTRNGQEPSRSLMGDARNAGDEQYIQAGLGQRVNNLLSQFTNLISLGEKGIED
ILQNQRGTELKLATENKSGRESEEANVEKILEVSNPQDMFKNFRLQNDLDSVQSPFRLPDADLSRELDSA
SFKDALDLKLPGNGEREIDLALEKVKVGETETSDLKVGQDESFVPAQLMKVETPEEKDDIIEQMVLRIRQ
DGETDENTVSGPGVAESLDIEAKGESAIAS1
```

```
;  TOIG of: aar42456  check: 3609  from: 1  to: 543
;
; ID   AAR42456 standard; protein; 543 AA.
; XX
; AC   AAR42456;
; XX
; DT   25-MAR-2003  (revised)
; DT   27-MAY-1994  (first entry)
; XX
; DE   Enzyme involved in eicosapentaenoic acid (EPA) synthesis.
; XX
; KW   EPA; eicosapentaenoic acid synthetase; drug; anticoagulant; hypolipemic;
; KW   hypoglycemic; antihypertensive; anticancer; pesticide; foodstuff;
; KW   additive.
; XX
; OS   Shewanella putrefaciens.
; XX
; PN   WO9323545-A1.
; XX
; PD   25-NOV-1993.
; XX
; PF   14-MAY-1993;   93WO-JP000641.
; XX
; PR   15-MAY-1992;   92JP-00147945.
; XX
; PA   (SAGA ) SAGAMI CHEM RES CENTRE.
; XX
; PI   Yazawa K,  Yamada A,  Kato S,  Kondo K;
; XX
; DR   WPI; 1993-386577/48.
; DR   N-PSDB; AAQ51128.
; XX
; PT   Gene coding for eicosa-penta:enoic acid synthetase - is isolated from
; PT   Pseudomonas, Alteromonas or Shewanella and used for recombinant prodn. of
; PT   eicosa-penta:enoic acid.
; XX
; PS   Claim 10; Page 91-94; 106pp; Japanese.
; XX
; CC   EPA is useful as a drug, having anticoagulant, hypolipemic, hypoglycemic,
; CC   antihypertensive and anticancer activity. It is also a pesticide and is
; CC   useful as a nutritional foodstuff and animal feed additive. (Updated on
; CC   25-MAR-2003 to correct PN field.)
; XX
; SQ   Sequence 543 AA;
;
;  AAR42456  Length: 543  April 15, 2004 09:20  Type: P  Check: 3609  ..
aar42456
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: aar99465  check: 3609  from: 1  to: 543
;
; ID   AAR99465 standard; protein; 543 AA.
; XX
; AC   AAR99465;
; XX
; DT   30-JAN-1997  (first entry)
; XX
; DE   Biosynthetic enzyme of icosapentaenoic acid synthase.
; XX
; KW   Icosapentaenoic acid synthase; EPA; drugs; agrochemicals; foodstuffs;
; KW   animal feed; lipid balance correction; antihypertensive;
; KW   antiinflammatory; anticancer agent.
; XX
; OS   Shewanella putrefaciens.
; XX
; PN   WO9621735-A1.
; XX
; PD   18-JUL-1996.
; XX
; PF   12-JAN-1996;   96WO-JP000030.
; XX
; PR   13-JAN-1995;   95JP-00004299.
; XX
; PA   (SAGA ) SAGAMI CHEM RES CENTRE.
; XX
; PI   Yazawa K,  Yamada A,  Kondo K;
; XX
; DR   WPI; 1996-342288/34.
; DR   N-PSDB; AAT34137.
; XX
; PT   Production of icosapentaenoic acid using transformed E. coli - uses DNA
; PT   coding for icosapentaenoic acid synthase derived from Shewanella strain.
; XX
; PS   Claim 7; Page 128-131; 145pp; English.
; XX
; CC   The DNA sequence (AAT34137) which encodes the biosynthetic enzymes of
; CC   icosapentaenoic acid (EPA) can be used to transform Escherichia coli. The
; CC   DNA sequence allows efficient microbial production of EPA, which is a raw
; CC   material for drugs, agrochemicals, foods and animal feedstuffs. EPA is
; CC   also useful for lipid balance correction and as an antihypertensive,
; CC   antiinflammatory and anticancer agent
; XX
; SQ   Sequence 543 AA;
;
;  AAR99465  Length: 543  April 15, 2004 09:20  Type: P  Check: 3609  ..
aar99465
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: aau65346  check: 8069  from: 1  to: 79
;
; ID   AAU65346 standard; protein; 79 AA.
; XX
; AC   AAU65346;
; XX
; DT   27-FEB-2002  (first entry)
; XX
; DE   Propionibacterium acnes immunogenic protein #26242.
; XX
; KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
; KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
; KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
; KW   dermatological; osteopathic; neuroprotectant.
; XX
; OS   Propionibacterium acnes.
; XX
; PN   WO200181581-A2.
; XX
; PD   01-NOV-2001.
; XX
; PF   20-APR-2001; 2001WO-US012865.
; XX
; PR   21-APR-2000; 2000US-0199047P.
; PR   02-JUN-2000; 2000US-0208841P.
; PR   07-JUL-2000; 2000US-0216747P.
; XX
; PA   (CORI-) CORIXA CORP.
; XX
; PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
; PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
; XX
; DR   WPI; 2001-616774/71.
; DR   N-PSDB; AAS59666.
; XX
; PT   Propionibacterium acnes polypeptides and nucleic acids useful for
; PT   vaccinating against and diagnosing infections, especially useful for
; PT   treating acne vulgaris.
; XX
; PS   Example 1; SEQ ID NO 26541; 1069pp; English.
; XX
; CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
; CC   polypeptides. The proteins and their associated DNA sequences are used in
; CC   the treatment, prevention and diagnosis of medical conditions caused by
; CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
; CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
; CC   P. acnes is also involved in infections of bone, joints and the central
; CC   nervous system, however it is particularly involved in the inflammatory
; CC   lesions associated with acne vulgaris. A method for detecting the
; CC   presence or absence of P. acnes in a patient comprises contacting a
; CC   sample with a binding agent that binds to the proteins of the invention
; CC   and determining the amount of bound protein in the sample. The
; CC   polypeptides may be used as antigens in the production of antibodies
; CC   specific for P. acnes proteins. These antibodies can be used to
; CC   downregulate expression and activity of P. acnes polypeptides and
; CC   therefore treat P. acnes infections. The antibodies may also be used as
; CC   diagnostic agents for determining P. acnes presence, for example, by
; CC   enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
; CC   this patent did not form part of the printed specification, but was
; CC   obtained in electronic format directly from WIPO at
; CC   ftp.wipo.int/pub/published_pct_sequences
; XX
; SQ   Sequence 79 AA;
;
;   AAU65346  Length: 79  April 15, 2004 09:23  Type: P  Check: 8069  ..
aau65346
TIEVSPGVVMARAPWAAPNSTAFCRSPVSIRPATRPAAKESPPPTRSRISRLGRSTERTKPVSFDHAIAD
QSLRLAVRT1
```

```
;  TOIG of: aau66759  check: 99  from: 1  to: 113
;
; ID   AAU66759 standard; protein; 113 AA.
; XX
; AC   AAU66759;
; XX
; DT   13-FEB-2002  (first entry)
; XX
; DE   Propionibacterium acnes immunogenic protein #27655.
; XX
; KW   SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
; KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
; KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
; KW   dermatological; osteopathic; neuroprotectant.
; XX
; OS   Propionibacterium acnes.
; XX
; PN   WO200181581-A2.
; XX
; PD   01-NOV-2001.
; XX
; PF   20-APR-2001; 2001WO-US012865.
; XX
; PR   21-APR-2000; 2000US-0199047P.
; PR   02-JUN-2000; 2000US-0208841P.
; PR   07-JUL-2000; 2000US-0216747P.
; XX
; PA   (CORI-) CORIXA CORP.
; XX
; PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
; PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
; XX
; DR   WPI; 2001-616774/71.
; DR   N-PSDB; AAS59754.
; XX
; PT   Propionibacterium acnes polypeptides and nucleic acids useful for
; PT   vaccinating against and diagnosing infections, especially useful for
; PT   treating acne vulgaris.
; XX
; PS   Example 1; SEQ ID NO 27954; 1069pp; English.
; XX
; CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
; CC   polypeptides. The proteins and their associated DNA sequences are used in
; CC   the treatment, prevention and diagnosis of medical conditions caused by
; CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
; CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
; CC   P. acnes is also involved in infections of bone, joints and the central
; CC   nervous system, however it is particularly involved in the inflammatory
; CC   lesions associated with acne vulgaris. A method for detecting the
; CC   presence or absence of P. acnes in a patient comprises contacting a
; CC   sample with a binding agent that binds to the proteins of the invention
; CC   and determining the amount of bound protein in the sample. The
; CC   polypeptides may be used as antigens in the production of antibodies
; CC   specific for P. acnes proteins. These antibodies can be used to
; CC   downregulate expression and activity of P. acnes polypeptides and
; CC   therefore treat P. acnes infections. The antibodies may also be used as
; CC   diagnostic agents for determining P. acnes presence, for example, by
; CC   enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
; CC   this patent did not form part of the printed specification, but was
; CC   obtained in electronic format directly from WIPO at
; CC   ftp.wipo.int/pub/published_pct_sequences
; XX
; SQ   Sequence 113 AA;
;
;   AAU66759  Length: 113  April 15, 2004 09:24  Type: P  Check: 99  ..
aau66759
PERTANGSPSCHGSDQRFTRADPPAANDSTSQRVTIEVSPGVVMARAPWAAPNSTAFCRSPVSIRPATRP
AAKESPPPTRSRISRLGRSTERTKPVSFDHAIADQSLRLAVRT1
```

```
;  TOIG of: aaw37053  check: 3609  from: 1  to: 543
;
; ID   AAW37053 standard; protein; 543 AA.
; XX
; AC   AAW37053;
; XX
; DT   03-JUL-1998  (first entry)
; XX
; DE   S. putrefaciens EPO biosynthesis gene cluster ORF9 product.
; XX
; KW   SCRC-2874; FERM BP-1625; eicosapentaenoic acid; EPA;
; KW   biosynthesis gene cluster; synthetase.
; XX
; OS   Shewanella putrefaciens.
; XX
; PN   WO9801565-A1.
; XX
; PD   15-JAN-1998.
; XX
; PF   09-JUL-1997;   97WO-JP002371.
; XX
; PR   10-JUL-1996;   96JP-00180845.
; XX
; PA   (SAGA ) SAGAMI CHEM RES CENTRE.
; XX
; PI   Yazawa K,  Yamada A,  Kondo K,  Kato S;
; XX
; DR   WPI; 1998-101060/09.
; DR   N-PSDB; AAV00503.
; XX
; PT   Eicosapentaenoic acid produced by culture of transformed Escherichia coli
; PT   - containing an eicosapentaenoic acid synthetase gene derived from the
; PT   marine microorganism Shewanella.
; XX
; PS   Example 1; Page 93-97; 110pp; Japanese.
; XX
; CC   The present sequence is encoded by the Shewanella putrefaciens SCRC-2874
; CC   (FERM BP-1625) eicosapentaenoic acid (EPA) biosynthesis gene cluster. A
; CC   novel EPA (useful in drugs, pesticides, foods and feedstuffs) is encoded
; CC   by synthetase enzyme gene sequences comprising parts of the full sequence
; CC   of the synthetase gene from the marine microorganism S. putrefaciens SCRC
; CC   -2874 (FERM BP-1625), in which at least 1 of the 9 open reading frames
; CC   (ORF) (numbered 2-10) in the gene have been deleted. In particular the
; CC   gene sequences comprising the following parts of the full gene: (1) bases
; CC   8081-9441, 12314-13084 and 13889-32520; (2) bases 8081-9441, 12314-13084,
; CC   13889-32520 and 34627-35559; (3) bases 8081-9441, 12314-13084 and 13889-
; CC   35559; (4) bases 8081-9441, 9681-13084 and 13889-32520; (5) bases 8081-
; CC   9441, 9681-13084, 13889-32520 and 34627-35564; and (6) bases 8081-9441,
; CC   9681-13084 and 13889-35564, are claimed
; XX
; SQ   Sequence 543 AA;
;
;  AAW37053  Length: 543  April 15, 2004 09:21  Type: P  Check: 3609  ..
aaw37053
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: aaw89403  check: 3609  from: 1  to: 543
;
; ID   AAW89403 standard; protein; 543 AA.
; XX
; AC   AAW89403;
; XX
; DT   19-OCT-1999  (first entry)
; XX
; DE   S. putrefaciens PKS-like cluster ORF 9 protein.
; XX
; KW   Polyketide-like synthesis; PKS; PKS-like gene; PUFA; DHA; transgenic;
; KW   poly-unsaturated fatty acid; eicosapentenoic acid; docosahexanoic acid;
; KW   EPA; oil; dietary supplement; infant feeeding formulation; malnutrition;
; KW   intravenous feeding formulation; cooking oil; fat; anti-inflammatory;
; KW   cholesterol; open reading frame; ORF.
; XX
; OS   Shewanella putrefaciens.
; XX
; PN   WO9855625-A1.
; XX
; PD   10-DEC-1998.
; XX
; PF   04-JUN-1998;   98WO-US011639.
; XX
; PR   04-JUN-1997;   97US-0048650P.
; XX
; PA   (CALJ ) CALGENE LLC.
; XX
; PI   Facciotti D,  Metz JG,  Lassner M;
; XX
; DR   WPI; 1999-070271/06.
; DR   N-PSDB; AAZ00331.
; XX
; PT   New nucleic acid encoding polyketide-like synthesis enzymes of Vibrio
; PT   marinus - and transformed plants and microbes that produce
; PT   polyunsaturated fatty acids, useful as pharmaceuticals and food
; PT   supplements.
; XX
; PS   Claim 17; Fig 4; 153pp; English.
; XX
; CC   The invention provides polyketide-like synthesis (PKS)-like genes that
; CC   are used for the production of long chain poly-unsaturated fatty acid
; CC   (PUFA) productions. Genes responsible for eicosapentenoic acid (EPA)
; CC   production in Shewanella putrefaciens and novel genes associated with the
; CC   production of docosahexanoic acid (DHA) in Vibrio marinus are used to
; CC   generate transgenic plants that can express transgenes encoding PKS-like
; CC   genes associated with PUFA production. The PKS-like genes are used to
; CC   transform plants and microbial cells to give recombinants having altered
; CC   contents of PUFA (specifically DHA and EPA). Oils from these plants are
; CC   useful as dietary supplements (in infant feeeding formulations, to give a
; CC   PUFA profile closer to that of human milk; for treating malnutrition; in
; CC   intravenous feeding formulations; in cooking oils, fats etc.), also as
; CC   anti-inflammatory agents and for reducing cholesterol levels. Fragments
; CC   from the genes are useful as probes to isolate related molecules or to
; CC   detect organisms that express PKS-like genes. The method facilitates
; CC   large scale production of PUFA by providing new pathways for their
; CC   synthesis or suppressing interfering pathways. Expression of PUFA in
; CC   seeds allows simple recovery, as oil which can be engineered to have a
; CC   particular PUFA profile. Expression in microbes also allows simple
; CC   recovery and control of PUFA profile and is not subject to external
; CC   variables such as weather or food supply. Sequences AAW89396-404
; CC   represent different ORF proteins of S. putrefaciens PKS-like cluster
; XX
; SQ   Sequence 543 AA;
;
;  AAW89403  Length: 543  April 15, 2004 09:21  Type: P  Check: 3609  ..
aaw89403
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
```

TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1

; TOIG of: abm61865 check: 8069 from: 1 to: 79
;
; ID ABM61865 standard; protein; 79 AA.
; XX
; AC ABM61865;
; XX
; DT 20-OCT-2003 (first entry)
; XX
; DE Propionibacterium acnes predicted ORF-encoded polypeptide #26541.
; XX
; KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
; KW immunostimulant; immune response; vaccine.
; XX
; OS Propionibacterium acnes.
; XX
; PN WO2003033515-A1.
; XX
; PD 24-APR-2003.
; XX
; PF 11-OCT-2002; 2002WO-US032727.
; XX
; PR 15-OCT-2001; 2001US-00978825.
; XX
; PA (CORI-) CORIXA CORP.
; XX
; PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
; PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
; PI Barth B, Vallieve-Douglass J;
; XX
; DR WPI; 2003-381789/36.
; DR N-PSDB; ACF64595.
; XX
; PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
; PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
; PT or for stimulating an immune response specific for a P. acnes protein.
; XX
; PS Example 1; SEQ ID NO 26541; 1481pp; English.
; XX
; CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
; CC encoding a Propionibacterium acnes protein. The invention also relates to
; CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
; CC immunogenic fragments of P. acnes polypeptides. The invention
; CC additionally encompasses expression vectors and host cells comprising a
; CC polynucleotide of the invention; antibodies against polypeptides of the
; CC invention; fusion proteins comprising a polypeptide of the invention; a
; CC method for stimulating an immune response specific for a P. acnes
; CC polypeptide and an isolated T cell population comprising T cells prepared
; CC via this method; a vaccine composition (comprising P. acnes polypeptides,
; CC polynucleotides, antibodies, fusion proteins, T cell populations, or
; CC antigen-presenting cells that express the polypeptide); a method and kit
; CC for detecting or determining the presence or absence of P. acnes in a
; CC patient; and a method for inhibiting the development of P. acnes in a
; CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
; CC proteins, T cell populations or antigen-presenting cells that express the
; CC polypeptides are useful for diagnosing, preventing or treating acne
; CC vulgaris, or for stimulating an immune response specific for a P. acnes
; CC protein. The polynucleotides can also be used as probes or primers for
; CC nucleic acid hybridisation. The vaccine composition is useful for the
; CC stimulation of an immune response against P. acnes, or for treating acne,
; CC and the kit is useful for performing a diagnostic assay. The present
; CC sequence represents a polypeptide predicted to be encoded by an ORF (open
; CC reading frame) contained within the P. acnes polynucleotides of the
; CC invention. Note: The sequence data for this patent did not form part of
; CC the printed specification, but was obtained in electronic format directly
; CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
; XX
; SQ Sequence 79 AA;
;
; ABM61865 Length: 79 April 15, 2004 09:24 Type: P Check: 8069 ..
abm61865
TIEVSPGVVMARAPWAAPNSTAFCRSPVSIRPATRPAAKESPPPTRSRISRLGRSTERTKPVSFDHAIAD
QSLRLAVRT1

```
;  TOIG of: abm63278  check: 99  from: 1  to: 113
;
; ID   ABM63278 standard; protein; 113 AA.
; XX
; AC   ABM63278;
; XX
; DT   20-OCT-2003  (first entry)
; XX
; DE   Propionibacterium acnes predicted ORF-encoded polypeptide #27954.
; XX
; KW   Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
; KW   immunostimulant; immune response; vaccine.
; XX
; OS   Propionibacterium acnes.
; XX
; PN   WO2003033515-A1.
; XX
; PD   24-APR-2003.
; XX
; PF   11-OCT-2002; 2002WO-US032727.
; XX
; PR   15-OCT-2001; 2001US-00978825.
; XX
; PA   (CORI-) CORIXA CORP.
; XX
; PI   Mitcham JL,  Skeiky YAW,  Persing DH,  Bhatia A,  Maisonneuve JL;
; PI   Zhang Y,  Wang S,  Jen S,  Lodes MJ,  Benson DR,  Jones R,  Carter D;
; PI   Barth B,  Vallieve-Douglass J;
; XX
; DR   WPI; 2003-381789/36.
; DR   N-PSDB; ACF64683.
; XX
; PT   New Propionibacterium acnes polypeptides and polynucleotides encoding the
; PT   polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
; PT   or for stimulating an immune response specific for a P. acnes protein.
; XX
; PS   Example 1; SEQ ID NO 27954; 1481pp; English.
; XX
; CC   The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
; CC   encoding a Propionibacterium acnes protein. The invention also relates to
; CC   polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
; CC   immunogenic fragments of P. acnes polypeptides. The invention
; CC   additionally encompasses expression vectors and host cells comprising a
; CC   polynucleotide of the invention; antibodies against polypeptides of the
; CC   invention; fusion proteins comprising a polypeptide of the invention; a
; CC   method for stimulating an immune response specific for a P. acnes
; CC   polypeptide and an isolated T cell population comprising T cells prepared
; CC   via this method; a vaccine composition (comprising P. acnes polypeptides,
; CC   polynucleotides, antibodies, fusion proteins, T cell populations, or
; CC   antigen-presenting cells that express the polypeptide); a method and kit
; CC   for detecting or determining the presence or absence of P. acnes in a
; CC   patient; and a method for inhibiting the development of P. acnes in a
; CC   patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
; CC   proteins, T cell populations or antigen-presenting cells that express the
; CC   polypeptides are useful for diagnosing, preventing or treating acne
; CC   vulgaris, or for stimulating an immune response specific for a P. acnes
; CC   protein. The polynucleotides can also be used as probes or primers for
; CC   nucleic acid hybridisation. The vaccine composition is useful for the
; CC   stimulation of an immune response against P. acnes, or for treating acne,
; CC   and the kit is useful for performing a diagnostic assay. The present
; CC   sequence represents a polypeptide predicted to be encoded by an ORF (open
; CC   reading frame) contained within the P. acnes polynucleotides of the
; CC   invention. Note: The sequence data for this patent did not form part of
; CC   the printed specification, but was obtained in electronic format directly
; CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
; XX
; SQ   Sequence 113 AA;
;
;   ABM63278  Length: 113  April 15, 2004 09:24  Type: P  Check: 99  ..
abm63278
PERTANGSPSCHGSDQRFTRADPPAANDSTSQRVTIEVSPGVVMARAPWAAPNSTAFCRSPVSIRPATRP
AAKESPPPTRSRISRLGRSTERTKPVSFDHAIADQSLRLAVRT1
```

```
! FINDPATTERNS on pir:* allowing 0 mismatches

!        1 E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R,I)(T,S,N)R

      (4)  XPBE12  ck: 1901  len: 871   ! major antigenic structural protein p100 - h

1                  E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                     E(S)(Q)(F)(S)(D)(I)(N)R(A)
           121: NKEKF   ESQFSDINRA   LLRLG

      (2)  D97264  ck: 8688  len: 497   ! galactose-1-phosphate uridyltransferase [im

1                  E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                     E(H)(I)(P)(M)(K)(I)(S)R(I)
           226: IILNN   EHIPMKISRI   TFENL

      (3)  T30186  ck: 3609  len: 543   ! hypothetical protein 8 - Shewanella sp.

1                  E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                     E(Q)(Q)(L)(K)(D)(F)(S)R(A)
            30: DVQVM   EQQLKDFSRA   CYVVN

      (1)  B90396  ck: 6793  len: 220   ! hypothetical protein SSO2259 [imported] - S

1                  E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                     E(E)(I)(I)(N)(D)(L)(S)R(I)
           137: LINHW   EEIINDLSRI   DLTNF


 Databases searched:
       NBRF, Release 78.0, Released on 24Nov2003, Formatted on 25Nov2003

     Total finds:              4
    Total length:     96,191,526
 Total sequences:        283,366
        CPU time:       02:57.60
```

```
;  TOIG of: b90396  check: 6793  from: 1  to: 220
;
; P1;B90396 - hypothetical protein SSO2259 [imported] - Sulfolobus solfataricus
; C;Species: Sulfolobus solfataricus
; C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
; C;Accession: B90396
; R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez,
;  M.J.; Chan-Weiher, C.C.Y.; Clausen, I.G.; Curtis, B.A.; De Moors, A.; Erauso,
;  G.; Fletcher, C.; Gordon, P.M.K.; Heikamp-de Jong, I.; Jeffries, A.C.; Kozera,
;  C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.; Schenk, M.E.;
;  Theriault, C.; Tolstrup, N.; Charlebois, R.L.; Doolittle, W.F.; Duguet, M.;
;  Gaasterland, T.; Garrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
; submitted to GenBank, April 2001
; A;Description: Sulfolobus solfataricus complete genome.
; A;Reference number: A99139
; A;Accession: B90396
; A;Status: preliminary
; A;Molecule type: DNA
; A;Residues: 1-220 <KUR>
; A;Cross-references: GB:AE006641; NID:g13815561; PIDN:AAK42425.1; GSPDB:GN00155
; C;Genetics:
; A;Gene: SSO2259
;
;   B90396  Length: 220  April 15, 2004 09:27  Type: P  Check: 6793  ..
b90396
MTNGEFYALDNLLKDEKLNSLNKILDVINTTDRLGILDVIKGILEDENTIGKIIGSLTSDDVLELLVNWD
KVIKTTKLFINEDNIYNIQFLINLIDKVRSKGILDPIIGLLEDEESLGKIINALINDFTLNLINHWEII
NDLSRIDLTNFKYYTLLVSATGEALKTENVKPITSIWEIYKLLKDPDIQRGLGVAASVLKRIGKLYVPDK
GLAFEVEKKL1
```

```
;  TOIG of: d97264  check: 8688  from: 1  to: 497
;
; P1;D97264 - galactose-1-phosphate uridyltransferase [imported] - Clostridium
;  acetobutylicum
; C;Species: Clostridium acetobutylicum
; C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 27-Oct-2003
; C;Accession: D97264
; R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson,
;  R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hitti, J.; Wolf, Y.I.; Tatusov, R.L.;
;  Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.;
;  Koonin, E.V.; Smith, D.R.
; J. Bacteriol. 183, 4823-4838, 2001
; A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing
;  Bacterium Clostridium acetobutylicum.
; A;Reference number: A96900; MUID:21359325; PMID:21359325
; A;Accession: D97264
; A;Status: preliminary
; A;Molecule type: DNA
; A;Residues: 1-497 <KUR>
; A;Cross-references: GB:AE001437; PIDN:AAK80903.1; PID:g15026015; GSPDB:GN00168
; A;Experimental source: Clostridium acetobutylicum ATCC824
; C;Genetics:
; A;Gene: CAC2961
; C;Superfamily: galactose-1-phosphate uridylyltransferase, Bacillus type
;
;   D97264  Length: 497  April 15, 2004 09:26  Type: P  Check: 8688  ..
d97264
MINHEINKLLAFSLKKGLIQEDDKIYSSNMLAGLFNLDNFYFEEISDVPSTATAILNQLLAYAVKENLIN
DTVAERDLFDTKIMNCVMPRPSEVINNFNRLLNNSPKEATSYYYKLSIASNYIRKDRIDKNITWKTPTEY
GDLDITINLSKPEKDPRDIAKAKLSKSTSYPKCLLCKENEGFYGNINHPARQTLRIIPLELNKSKWFLQY
SPYTYYNEHCIILNNEHIPMKISRITFFENLLSFIDILPHYFAGSNADLPIVGGSILSHDHYQGGRYTFAM
EKAPVEKEYSIKGYEDISVGRVKWPMSVIRISSKNKTKLINLAEHILTSWRNYSDKTQSILSHTGSEPHN
TITPIARKRNEEYELDLVLRNNRTDENYPLGIFHPHNEVHHIKKENIGLIEVMGLAVLPARLKSELALIK
ENLIEKKKDISNDSTISKHNTWYKYILDNYKNISEENIDCILKKEVGIKFLEVLKHAGVFKRNSSGLSAF
DKFINIL1
```

```
;  TOIG of: t30186  check: 3609  from: 1  to: 543
;
; P1;T30186 - hypothetical protein 8 - Shewanella sp.
; C;Species: Shewanella sp.
; C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
; C;Accession: T30186
; R;Takeyama, H.; Takeda, D.; Yazawa, K.; Yamada, A.; Matsunaga, T.
; Microbiology 143, 2725-2731, 1997
; A;Title: Expression of the eicosapentaenoic acid synthesis gene cluster from
;  Shewanella sp. in a transgenic marine cyanobacterium, Synechococcus sp.
; A;Reference number: Z20764; MUID:97419510; PMID:9274025
; A;Accession: T30186
; A;Status: preliminary; translated from GB/EMBL/DDBJ
; A;Molecule type: DNA
; A;Residues: 1-543 <TAK>
; A;Cross-references: EMBL:U73935; NID:g2529413; PID:g2529421; PIDN:AAB81126.1
; A;Experimental source: strain SCRC-2738
;
;   T30186  Length: 543  April 15, 2004 09:26  Type: P  Check: 3609  ..
t30186
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: xpbe12  check: 1901  from: 1  to: 871
;
; P1;XPBE12 - major antigenic structural protein p100 - human herpesvirus 6
;  (strain U1102)
; C;Species: human herpesvirus 6
; C;Date: 30-Jun-1993 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
; C;Accession: T09303; A42533
; R;Nicholas, J.; Martin, M.
; J. Virol. 68, 597-610, 1994
; A;Title: Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
;  genome of human herpesvirus 6 encoding human cytomegalovirus immediate-early
;  gene homologs and transactivating functions.
; A;Reference number: Z16644; MUID:94118404; PMID:8289364
; A;Accession: T09303
; A;Status: preliminary; translated from GB/EMBL/DDBJ
; A;Molecule type: DNA
; A;Residues: 1-871 <NIC>
; A;Cross-references: EMBL:L25528; NID:g451932; PIDN:AAA16716.1; PID:g451934
; R;Neipel, F.; Ellinger, K.; Fleckenstein, B.
; J. Virol. 66, 3918-3924, 1992
; A;Title: Gene for the major antigenic structural protein (p100) of human
;  herpesvirus 6.
; A;Reference number: A42533; MUID:92260671; PMID:1374813
; A;Accession: A42533
; A;Molecule type: DNA
; A;Residues: 2-871 <NEI>
; A;Cross-references: GB:M87287; NID:g330673; PIDN:AAA46012.1; PID:g330674
; C;Genetics:
; A;Gene: P1LF1
; C;Superfamily: human herpesvirus large structural phosphoprotein; large
;  structural phosphoprotein homology
; C;Keywords: phosphoprotein
; F;7-368/Domain: large structural phosphoprotein homology <CLS>
;
;   XPBE12  Length: 871  April 15, 2004 09:26  Type: P  Check: 1901  ..
xpbe12
NMDLQRHPIPFAWLDRDKVERLTDFLSNLERLDNVDLREHPHVTNSCVVREGDDVDDLKTLYNLLVLWLM
YHYVLSKRKPDYNAIWQDITKLQSVVNEYLNSKGLNKGIFENMFTNKEKFESQFSDINRALLRLGNFIKW
GSNVAIDTPYVNLTAEDSSEIENNLQDAEKNMLWYTVYNINDPWDENGYLITSINKLIYLGKLFLALTQS
WSKLEKVAMSQIVITQNHLSGHLRRHDNFNIVYSHRVLQTPLTGQRVESFLKIITSDYDIIKSSLESHSA
SKAFSMSEIGPNSLMDFVPLRGDIHSNLTLPSMSIDTKKSSLDPARLKKSNSRSLDSFLRMQRQPKFLEL
DSVDNAGEKILLKEATLGGENVKATTPASSVSLMSGVESPSSFTSTNLDLPLSSFTSTNLDLRDKSHGNY
KIGPSGILDFNVKFPPNAQLNTNGVDLLQDKTSIGSPSSGITDVVNGFANLNLHQNKSNVSPPWSRNTAA
NADFLDPVHRFVPEQTGTPFVLNNSDVAGSEAKHTTYSTETGVSPRNVFLIKDLRGKDGFRKQKQSDIPK
SLTKERNDKAIMHSREVTGDSGDATETVGARNSPALRKIKQANDFFAGLNKKNDRDVLRGGKGNSKDLHS
GGNAKKKEMSGKFNDDKEMTRNGQEPSRSLMGDARNAGDEQYIQAGLGQRVNNLLSQFTNLISLGEKGIE
DILQNQRGTELKLATENKSGRESEEANVEKILEVSNPQDMFKNFRLQNDLDSVQSPFRLPDADLSRELDS
ASFKDALDLKLPGNGEREIDLALEKVKVGETETSDLKVGQDESFVPAQLMKVETPEEKDDIIEQMVLRIR
QDGETDENTVSGPGVAESLDIEAKGESAIAS1
```

```
! FINDPATTERNS on swp:* allowing 0 mismatches
!       1 E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R,I)(T,S,N)R

(1)  GALT_CLOAB  ck: 8688  len: 497   ! Q97ez4 clostridium acetobutylicum. galactos
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)-(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(H)(I)(P)(M)(K)(I)(S)R(I)
         226: IILNN   EHIPMKISRI   TFENL

(2)  P100_HSV6U  ck: 500   len: 870   ! Q00701 human herpesvirus (type 6 / strain u
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(S)(Q)(F)(S)(D)(I)(N)R(A)
         120: NKEKF   ESQFSDINRA   LLRLG

(3)  PFDA_METKA  ck: 9337  len: 157   ! Q8tuy7 methanopyrus kandleri. prefoldin alp
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(S)(S)(I)(S)(E)(L)(N)R(V)
          39: QIDLI   ESSISELNRV   EETLK

(4)    O33907  ck: 3609  len: 543   ! O33907 shewanella sp. scrc-2738. hypothetic
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(Q)(Q)(L)(K)(D)(F)(S)R(A)
          30: DVQVM   EQQLKDFSRA   CYVVN

(7)  Q9ESP3  ck: 4988  len: 1,851 ! Q9esp3 rattus norvegicus (rat). mucin (frag
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(S)(N)(S)(T)(T)(L)(S)R(V)
         130: LRRGL   ESNSTTLSRV   IMKLD

(5)  Q80Z21  ck: 6085  len: 1,726 ! Q80z21 mus musculus (mouse). secreted gel-f
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(S)(N)(T)(T)(T)(L)(S)R(V)
         126: LRRVQ   ESNTTTLSRV   TMKLD

(6)  Q97WG2  ck: 6793  len: 220   ! Q97wg2 sulfolobus solfataricus. hypothetica
1                E(Q,S,H,Y,E)(S,Q,I,T,N,P)(F,T,S,P,L,I)(N,S,K,M,T,P)(D,K,T,E)(F,L,R
                   E(E)(I)(I)(N)(D)(L)(S)R(I)
         137: LINHW   EEIINDLSRI   DLTNF

Databases searched:
      SWISS-PROT, Release 42.7, Released on 15Dec2003, Formatted on 15Dec2003
      SPTREMBL, Release 25.0, Released on 17Oct2003, Formatted on 18Oct2003

    Total finds:           7
   Total length:     367,588,357
Total sequences:       1,158,722
       CPU time:         12:01.74
```

```
;  TOIG of: galt_cloab  check: 8688  from: 1  to: 497
;
; ID   GALT_CLOAB      STANDARD;      PRT;   497 AA.
; AC   Q97EZ4;
; DT   28-FEB-2003 (Rel. 41, Created)
; DT   28-FEB-2003 (Rel. 41, Last sequence update)
; DT   10-OCT-2003 (Rel. 42, Last annotation update)
; DE   Galactose-1-phosphate uridylyltransferase (EC 2.7.7.12) (Gal-1-P
; DE   uridylyltransferase) (UDP-glucose--hexose-1-phosphate
; DE   uridylyltransferase).
; GN   GALT OR CAC2961.
; OS   Clostridium acetobutylicum.
; OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
; OC   Clostridium.
; OX   NCBI_TaxID=1488;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
; RX   MEDLINE=21359325; PubMed=11466286;
; RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
; RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
; RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
; RA   Bennett G.N., Koonin E.V., Smith D.R.;
; RT   "Genome sequence and comparative analysis of the solvent-producing
; RT   bacterium Clostridium acetobutylicum.";
; RL   J. Bacteriol. 183:4823-4838(2001).
; CC   -!- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
; CC       alpha-D-glucose 1-phosphate + UDP-galactose.
; CC   -!- PATHWAY: Galactose metabolism; second step.
; CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
; CC   -!- SIMILARITY: Belongs to the galactose-1-phosphate
; CC       uridylyltransferase family 2.
; CC   ---------------------------------------------------------------------
; CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
; CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
; CC   the European Bioinformatics Institute. There are no restrictions on its
; CC   use by non-profit -institutions as long as its content is in no way
; CC   modified and this statement is not removed. Usage by and for commercial
; CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
; CC   or send an email to license@isb-sib.ch).
; CC   ---------------------------------------------------------------------
; DR   EMBL; AE007793; AAK80903.1; -.
; DR   PIR; D97264; D97264.
; DR   HAMAP; MF_00571; -; 1.
; DR   InterPro; IPR000766; GalP_transf_II.
; DR   InterPro; IPR005850; GalP_Utransf_C.
; DR   InterPro; IPR005849; GalP_Utransf_N.
; DR   InterPro; IPR005934; GalT2.
; DR   Pfam; PF02744; GalP_UDP_tr_C; 1.
; DR   Pfam; PF01087; GalP_UDP_transf; 1.
; DR   TIGRFAMs; TIGR01239; galT_2; 1.
; DR   PROSITE; PS01163; GAL_P_UDP_TRANSF_II; 1.
; KW   Transferase; Nucleotidyltransferase; Galactose metabolism;
; KW   Complete proteome.
; SQ   SEQUENCE   497 AA;  57288 MW;  81E9D089CDA9F5CC CRC64;
;
; GALT_CLOAB  Length: 497  April 15, 2004 09:30  Type: P  Check: 8688  ..
galtcloab
MINHEINKLLAFSLKKGLIQEDDKIYSSNMLAGLFNLDNFYFEEISDVPSTATAILNQLLAYAVKENLIN
DTVAERDLFDTKIMNCVMPRPSEVINNFNRLLNNSPKEATSYYYKLSIASNYIRKDRIDKNITWKTPTEY
GDLDITINLSKPEKDPRDIAKAKLSKSTSYPKCLLCKENEGFYGNINHPARQTLRIIPLELNKSKWFLQY
SPYTYYNEHCIILNNEHIPMKISRITFENLLSFIDILPHYFAGSNADLPIVGGSILSHDHYQGGRYTFAM
EKAPVEKEYSIKGYEDISVGRVKWPMSVIRISSKNKTKLINLAEHILTSWRNYSDKTQSILSHTGSEPHN
TITPIARKRNEEYELDLVLRNNRTDENYPLGIFHPHNEVHHIKKENIGLIEVMGLAVLPARLKSELALIK
ENLIEKKKDISNDSTISKHNTWYKYILDNYKNISEENIDCILKKEVGIKFLEVLKHAGVFKRNSSGLSAF
DKFINIL1
```

```
;  TOIG of: p100_hsv6u  check: 500  from: 1  to: 870
;
; ID   P100_HSV6U      STANDARD;      PRT;   870 AA.
; AC   Q00701;
; DT   01-APR-1993 (Rel. 25, Created)
; DT   01-APR-1993 (Rel. 25, Last sequence update)
; DT   16-OCT-2001 (Rel. 40, Last annotation update)
; DE   Large structural phosphoprotein (PP100) (P100) (Major antigenic
; DE   structural protein).
; GN   U11 OR P1LF1.
; OS   Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
; OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
; OC   Betaherpesvirinae; Roseolovirus.
; OX   NCBI_TaxID=10370;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RX   MEDLINE=92260671; PubMed=1374813;
; RA   Neipel F., Ellinger K., Fleckenstein B.;
; RT   "Gene for the major antigenic structural protein (p100) of human
; RT   herpesvirus 6.";
; RL   J. Virol. 66:3918-3924(1992).
; RN   [2]
; RP   SEQUENCE FROM N.A.
; RX   MEDLINE=94118404; PubMed=8289364;
; RA   Nicholas J., Martin M.;
; RT   "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
; RT   genome of human herpesvirus 6 encoding human cytomegalovirus
; RT   immediate-early gene homologs and transactivating functions.";
; RL   J. Virol. 68:597-610(1994).
; RN   [3]
; RP   SEQUENCE FROM N.A.
; RX   MEDLINE=95266321; PubMed=7747482;
; RA   Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
; RA   Martin M.E., Efstathiou S., Craxton M., Macaulay H.A.;
; RT   "The DNA sequence of human herpesvirus-6: structure, coding content,
; RT   and genome evolution.";
; RL   Virology 209:29-51(1995).
; CC   -!- SIMILARITY: TO THE LARGE STRUCTURAL PHOSPHOPROTEINS OF HSV-7 AND
; CC       HCMV UL32.
; CC   ---------------------------------------------------------------------
; CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
; CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
; CC   the European Bioinformatics Institute. There are no restrictions on its
; CC   use by non-profit institutions as long as its content is in no way
; CC   modified and this statement is not removed. Usage by and for commercial
; CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
; CC   or send an email to license@isb-sib.ch).
; CC   ---------------------------------------------------------------------
; DR   EMBL; M87287; AAA46012.1; -.
; DR   EMBL; L25528; AAA16716.1; ALT_INIT.
; DR   EMBL; X83413; CAA58438.1; -.
; KW   Matrix protein; Phosphorylation.
; SQ   SEQUENCE   870 AA;  97071 MW;  F25954DEA19BF824 CRC64;
;
;  P100_HSV6U  Length: 870  April 15, 2004 09:31  Type: P  Check: 500  ..
p100hsv6u
MDLQRHPIPFAWLDRDKVERLTDFLSNLERLDNVDLREHPHVTNSCVVREGDDVDDLKTLYNLLVLWLMY
HYVLSKRKPDYNAIWQDITKLQSVVNEYLNSKGLNKGIFENMFTNKEKFESQPSDINRALLRLGNFIKWG
SNVAIDTPYVNLTAEDSSEIENNLQDAEKNMLWYTVYNINDPWDENGYLITSINKLIYLGKLFLALTQSW
SKLEKVAMSQIVITQNHLSGHLRRHDNFNIVYSHRVLQTPLTGQRVESFLKIITSDYDIIKSSLESHSAS
KAFSMSEIGPNSLMDFVPLRGDIHSNLTLPSMSIDTKKSSLDPARLKKSNSRSLDSFLRMQRQPKFLELD
SVDNAGEKILLKEATLGGENVKATTPASSVSLMSGVESPSSFTSTNLDLPLSSFTSTNLDLRDKSHGNYK
IGPSGILDFNVKFPPNAQLNTNGVDLLQDKTSIGSPSSGITDVVNGFANLNLHQNKSNVSPPWSRNTAAN
ADFLDPVHRFVPEQTGTPFVLNNSDVAGSEAKHTTYSTETGVSPRNVFLIKDLRGKDGFRKQKQSDIPKS
LTKERNDKAIMHSREVTGDSGDATETVGARNSPALRKIKQANDFFAGLNKKNDRDVLRGGKGNSKDLHSG
GNAKKKEMSGKFNDDKEMTRNGQEPSRSLMGDARNAGDEQYIQAGLGQRVNNLLSQFTNLISLGEKGIED
ILQNQRGTELKLATENKSGRESEEANVEKILEVSNPQDMFKNFRLQNDLDSVQSPFRLPDADLSRELDSA
SFKDALDLKLPGNGEREIDLALEKVKVGETETSDLKVGQDESFVPAQLMKVETPEEKDDIIEQMVLRIRQ
DGETDENTVSGPGVAESLDIEAKGESAIAS1
```

```
;  TOIG of: pfda_metka  check: 9337  from: 1  to: 157
;
; ID   PFDA_METKA      STANDARD;      PRT;   157 AA.
; AC   Q8TUY7;
; DT   28-FEB-2003 (Rel. 41, Created)
; DT   28-FEB-2003 (Rel. 41, Last sequence update)
; DT   28-FEB-2003 (Rel. 41, Last annotation update)
; DE   Prefoldin alpha subunit (GimC alpha subunit).
; GN   PFDA OR MK1614.
; OS   Methanopyrus kandleri.
; OC   Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
; OC   Methanopyrus.
; OX   NCBI_TaxID=2320;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=AV19 / DSM 6324 / JCM 9639;
; RX   MEDLINE=21927647; PubMed=11930014;
; RA   Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
; RA   Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
; RA   Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
; RA   Malykh A.G., Koonin E.V., Kozyavkin S.A.;
; RT   "The complete genome of hyperthermophile Methanopyrus kandleri AV19
; RT   and monophyly of archaeal methanogens.";
; RL   Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
; CC   -!- FUNCTION: Molecular chaperone capable of stabilizing a range of
; CC       proteins. Seems to fulfil an ATP-independent, HSP70-like function
; CC       in archaeal de novo protein folding (By similarity).
; CC   -!- SUBUNIT: Heterohexamer of two alpha and four beta subunits (By
; CC       similarity).
; CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
; CC   -!- SIMILARITY: Belongs to the prefoldin alpha subunit family.
; CC   ---------------------------------------------------------------------
; CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
; CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
; CC   the European Bioinformatics Institute. There are no restrictions on its
; CC   use by non-profit institutions as long as its content is in no way
; CC   modified and this statement is not removed. Usage by and for commercial
; CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
; CC   or send an email to license@isb-sib.ch).
; CC   ---------------------------------------------------------------------
; DR   EMBL; AE010451; AAM02827.1; -.
; DR   HAMAP; MF_00308; -; 1.
; DR   InterPro; IPR004127; DUF232.
; DR   Pfam; PF02996; Prefoldin; 1.
; KW   Chaperone; Complete proteome.
; SQ   SEQUENCE   157 AA;  17446 MW;  B20BDA8CC978DFA1 CRC64;
;
; PFDA_METKA  Length: 157  April 15, 2004 09:31  Type: P  Check: 9337  ..
pfdametka
MAEKKNEQEIQQELQRLIAEINRLQGQMEAINAQIDLIEQSISELNRVEETLKGVKELEGDEEVLVPVGA
QSFVRACVTDTERVIVGIGAGVAVERTIDEALESIDDQRQELEKARAEAQQKLQELAQELQEKQRKAQEL
AQQLEGAQRIAQQSGGG1
;  TOIG of: o33907  check: 3609  from: 1  to: 543
;
; ID   O33907      PRELIMINARY;      PRT;   543 AA.
; AC   O33907;
; DT   01-JAN-1998 (TrEMBLrel. 05, Created)
; DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
; DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
; DE   Hypothetical protein.
; OS   Shewanella sp. SCRC-2738.
; OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
; OC   Alteromonadaceae; Shewanella.
; OX   NCBI_TaxID=53560;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=SCRC-2738;
; RX   MEDLINE=97419510; PubMed=9274025;
; RA   Takeyama H., Takeda D., Yazawa K., Yamada A., Matsunaga T.;
; RT   "Expression of the eicosapentaenoic acid synthesis gene cluster from
; RT   Shewanella sp. in a transgenic marine cyanobacterium, Synechococcus
; RT   sp.";
; RL   Microbiology 143:0-0(0).
; DR   EMBL; U73935; AAB81126.1; -.
; DR   PIR; T30186; T30186.
; KW   Hypothetical protein.
; SQ   SEQUENCE   543 AA;  59378 MW;  7233F53635B794C7 CRC64;
;
;  O33907  Length: 543  April 15, 2004 09:31  Type: P  Check: 3609  ..
o33907
MNPTATNEMLSPWPWAVTESNISFDVQVMEQQLKDFSRACYVVNHADHGFGIAQTADIVTEQAANSTDLP
VSAFTPALGTESLGDNNFRRVHGVKYAYYAGAMANGISSEELVIALGQAGILCGSFGAAGLIPSRVEAAI
NRIQAALPNGPYMFNLIHSPSEPALERGSVELFLKHKVRTVEASAFLGLTPQIVYYRAAGLSRDAQGKVV
VGNKVIAKVSRTEVAEKFMMPAPAKMLQKLVDDGSITAEQMELAQLVPMADDITAEADSGGHTDNRPLVT
LLPTILALKEEIQAKYQYDTPIRVGCGGGVGTPDAALATFNMGAAYIVTGSINQACVEAGASDHTRKLLA
TTEMADVTMAPAADMFEMGVKLQVVKRGTLFPMRANKLYEIYTRYDSIEAIPLDEREKLEKQVFRSSLDE
IWAGTVAHFNERDPKQIERAEGNPKRKMALIFRWYLGLSSRWSNSGEVGREMDYQIWAGPALGAFNQWAK
GSYLDNYQDRNAVDLAKHLMYGAAYLNRINSLTAQGVKVPAQLLRWKPNQRMA1
```

```
;  TOIG of: q80z21  check: 6085  from: 1  to: 1726
;
; ID   Q80Z21      PRELIMINARY;      PRT;  1726 AA.
; AC   Q80Z21;
; DT   01-JUN-2003 (TrEMBLrel. 24, Created)
; DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
; DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
; DE   Secreted gel-forming mucin (Fragment).
; GN   MUC5AC.
; OS   Mus musculus (Mouse).
; OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
; OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
; OX   NCBI_TaxID=10090;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=C57BL/6J;
; RA   Escande F., Buisine M.P.;
; RT   "The mouse secreted gel-forming mucin gene cluster.";
; RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
; DR   EMBL; AJ511870; CAD54412.1; -.
; DR   GO; GO:0016020; C:membrane; IEA.
; DR   GO; GO:0005215; F:transporter activity; IEA.
; DR   GO; GO:0006810; P:transport; IEA.
; DR   InterPro; IPR001064; Crystallin.
; DR   InterPro; IPR005829; Sug_transporter.
; DR   InterPro; IPR002919; TIL_Cysrich.
; DR   InterPro; IPR006552; VC_out.
; DR   InterPro; IPR001846; VWF_D.
; DR   Pfam; PF01826; TIL; 2.
; DR   Pfam; PF00094; vwd; 3.
; DR   SMART; SM00215; VWC_out; 2.
; DR   SMART; SM00216; VWD; 3.
; DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
; DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
; FT   NON_TER    1726   1726
; SQ   SEQUENCE   1726 AA;  186394 MW;  B90FC8517631057A CRC64;
;
;   Q80Z21  Length: 1726  April 15, 2004 09:32  Type: P  Check: 6085  ..
q80z21
MGVGRRKLVPFWVLALALACSQCTGQAQQDSLKSYHEHRSDVPHPQGHVGTPLNRVTIIPPLKTIPVVRA
FNPGHTRRVCSTWGNFHYKTFDGQVFYFPGLCNYVFSAHCGDAYEDFNIQLRRVQESNTTTLSRVTMKLD
GLVVELTKSSVLVNNHPVQLPFSQSGVLIELSNGYLKVVARLGLLFVWNEDDSLLLELDTKYTNKTCGLC
GDFNGSPKSNEFLSNNVRLTPLEFGNLQKMDGPTEQCQDPLPVPQKNCSARSGICEMILKGELFSGCAAL
VDISSYVEACRQDVCLCESLDPSDCICHTLAEYSRQCAHAGGQPQDWRGPNLCSQTCPLNMQHQECGSPC
VDTCSNPQHSQVCEDHCIAGCFCPEGMVLDDINQMGCVPVSQCACLYNGTLYAPGTNYSTDCTNTCSGGQ
WSCQDIPCAGTCSVMGGSHMSTFDGRQYTVHGDCTYVLSKPCDSNAFTVLVELRKCGLTESETCLKTVTL
NLGGGQTEIMVKATGEVFVNQIYTQLPVSTANATFFRPSTFFIVGETNLGLQLEIQLSPIMQTSVRLKPG
LRGLTCGLCGNFNSMQADDFQTISGVVEGTAAAFFNTFKTQAACPNVKNIFQDPCSLSVENEKYAQHWCS
LLTNASGPFSQCHATVNPSTFFSNCMYDTCNCEKSEDCMCAALSSYVRACAAKGVLLSDWRDGICTKPTI
TCPKSMTYQYHISTCQPTCRALNEKDVTCHVSFIPVDGCTCPKGTFLDDLGKCVQATSCPCYYKGSTVPN
GESVQDSGAICTCTQGALTCIGGPAPTPVCDAPMIYFDCHNATPGDTGAGCQKSCHTLDMTCYSSECVPG
CVCPNGLVADGNGGCVVTEDCPCVHNEATYRPGETIQVGCNNCTCENRMWQCTDKPCLATCAVYGDGHYI
TFDGQRYSFNGDCEYTLLQDNCGGNGSSQDAFRVITENIPCGTTGTTCSKSIKIFLGNYELKLSDSKMEV
VQKDVGQEPPYFVHQMGNYLVVETDIGLVLLWDKKTSIFLRLSPEFKGRVCGLCGNFDDNAINDFTTRSQ
SVVSDMLEFGNSWKLSPSCPDVLVPKDPCTANPYRKSWAQKQCSIINSETFSACHAHVEPAKYYEACVND
ACACDSGGDCECFCTTVAAYAQACHEVGVCVSWRTPDICPLFCDYYNPEGQCEWHYQPCGAPCMRTCQNP
TGQCLQDLRGLEGCYPKCPPTAPIFDEGTMQCVSNCTVTFPCRVNGKLYRPGASVPSDKNCDSCICTESG
VRCTHNAGACVCTYNGQQFHPGEIIYHTTDGIGGCISAHCRANGTIERSVDTCNSTTPTPPTTFSFSTPP
VMTSMQPSSTHSSPTPSVGSSGASSKAASTTSSILSVKSPVTAPMTMSTSASAVTTSGCREECLWSPWMD
VSRPGRGIDSGDFDTLENLRAHGYPICQVPKAVECRAEASPGVPLPELQQHLECSTTVGLICYNSDQLSG
LCDNYQIKVQCCTPVSCPTSQTTHVISSSRTTNLDNTTSSVPVTSTEHPYSSTVTSGSSTHTPGLSPSSS
VPSSPTPASSTPAPVSSTTVKTTLPITSPTPEPTPAISSVSISTSGSTMPSSETTHECKQELCNWTNWLD
GSYPGSGRNSGDFDTFVNLRSKGYKFCEKPRNVECRAQFFPNTPLEELGQNVTCSREEGLICLNKNQLPP
MCYNYEIRIECCTVVNNCSTASVTTHPTSHGVSTKTETNWTTHVYS1
```

```
;  TOIG of: q97wg2  check: 6793  from: 1  to: 220
;
; ID   Q97WG2      PRELIMINARY;      PRT;   220 AA.
; AC   Q97WG2;
; DT   01-OCT-2001 (TrEMBLrel. 18, Created)
; DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
; DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
; DE   Hypothetical protein SSO2259.
; GN   SSO2259.
; OS   Sulfolobus solfataricus.
; OC   Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
; OC   Sulfolobus.
; OX   NCBI_TaxID=2287;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=ATCC 35092 / DSM 1617 / P2;
; RX   MEDLINE=21332296; PubMed=11427726;
; RA   She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
; RA   Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
; RA   De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
; RA   Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
; RA   Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
; RA   Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
; RA   Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
; RT   "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
; RL   Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
; DR   EMBL; AE006830; AAK42425.1; -.
; DR   PIR; B90396; B90396.
; KW   Hypothetical protein; Complete proteome.
; SQ   SEQUENCE   220 AA;  24900 MW;  ECC55875E903DF2B CRC64;
;
;   Q97WG2  Length: 220  April 15, 2004 09:32  Type: P  Check: 6793  ..
q97wg2
MTNGEFYALDNLLKDEKLNSLNKILDVINTTDRLGILDVIKGILEDENTIGKIIGSLTSDDVLELLVNWD
KVIKTTKLFINEDNIYNIQFLINLIDKVRSKGILDPIIGLLEDEESLGKIINALINDFTLNLINHWEEII
NDLSRIDLTNFKYYTLLVSATGEALKTENVKPITSIWEIYKLLKDPDIQRGLGVAASVLKRIGKLYVPDK
GLAFEVEKKL1
```

```
;  TOIG of: q9esp3  check: 4988  from: 1  to: 1851
;
; ID   Q9ESP3      PRELIMINARY;      PRT;  1851 AA.
; AC   Q9ESP3;
; DT   01-MAR-2001 (TrEMBLrel. 16, Created)
; DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
; DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
; DE   Mucin (Fragment).
; GN   MUC5AC.
; OS   Rattus norvegicus (Rat).
; OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
; OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
; OX   NCBI_TaxID=10116;
; RN   [1]
; RP   SEQUENCE FROM N.A.
; RC   STRAIN=Wistar; TISSUE=Stomach;
; RA   Oinuma T., Suganuma T.;
; RT   "Rat gastric mucin Muc5AC: Sequence of its 5'-region contains
; RT   conserved D-domains and two leucine zipper motifs.";
; RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
; DR   EMBL; AB042530; BAB17787.1; -.
; DR   GO; GO:0016020; C:membrane; IEA.
; DR   GO; GO:0005215; F:transporter activity; IEA.
; DR   GO; GO:0006810; P:transport; IEA.
; DR   InterPro; IPR001064; Crystallin.
; DR   InterPro; IPR005829; Sug_transporter.
; DR   InterPro; IPR002919; TIL_Cysrich.
; DR   InterPro; IPR001007; VWF_C.
; DR   InterPro; IPR001846; VWF_D.
; DR   Pfam; PF01826; TIL; 3.
; DR   Pfam; PF00094; vwd; 3.
; DR   SMART; SM00214; VWC; 3.
; DR   SMART; SM00216; VWD; 3.
; DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
; DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
; FT   NON_TER    1851   1851
; SQ   SEQUENCE   1851 AA;  199403 MW;  879DE5B54929C52B CRC64;
;
;   Q9ESP3  Length: 1851  April 15, 2004 09:31  Type: P  Check: 4988  ..
q9esp3
MLHSMGVGRRKLAPFWVLALALTFNQHTGQALEDTRKSHLEHYSDLSQPQGHVGTPLNRVTIIPPLKTIP
VVRAFNPAHTRRVCSTWGNFHYKTFDGQVFYFPGLCNYVFSEHCGAAYEDFNIQLRRGLESNSTTLSRVI
MKLDGLVVELTKSSVLVNNHPVQLPFSQSGVLIELSNGYLKVVARLGLVFMWNDDDSLLLELDTKYANKT
CGLCGDFNGSPESSEFLSHNVRLTPLEFGNFQKMDGPTEQCQDPLPVPQKNCSIRSSICEEILKGQLFSN
CAALVDISSYLEACQQDLCLCESSDPSNCICHTLAEYSRQCAHAGGQPQNWRGPNLCPQTCLLNMEYQEC
GSPCVDTCSNPQHSQVCEDHCVAGCFCPEGMVLDDSNQTGCVPVSQCACLYNGTLYAPGTSYSTDCTKCT
CSGGQWSCQEVPCSGTCSVMGGSHISTFDERQYTVHGDCSYVLCKPYDSNAFTVLAELRKCGLTESETCL
KTVTLNLGGGKTVITVKATGEVFVNQIYTQLPVSTANATFFRPSTFFIIGQTNLGLQLEIQLHPIMQVSV
RIAPEFRGLTSGLCGNFNSMQADDFQTISGVVEGTAAAFFNTFKTQAACPNVKNIFEDPCSLSVENEKYA
QHWCSQLTDANGPFSQCHATVNPSTFFSNCMFDTCNCEKSEDCLCAALSSYVRACAAKGVLLSDWREGIC
AKPTITCPKSMTYQYHISTCQPTCRSLSEEDVTCHVNFIPVDGCTCPKGTFLDDSGKCVQATSCPCYYKG
SPVPNGESVHDNGAICTCTQGALTCIGGPVLTPVCDAPMIYFDCRNATPGDTGAGCQKSCHTLDMTCYSS
ECVPGCVCPNGLVADGNGSCVVAEDCPCVHNEATYRPGETIQVGCNNCTCENRMWQCTDKPCLATCAVYG
DGHYITFDGQRYSFSGDCEYTLLQDNCGGNGSSQDAFRVVTENIPCGTTGTTCSKGIKIFLGSYELKLSD
SKMEVVQKGVGQEPPYFVHQMGNYLVVETDIGLVLLWDKKTSIFLRLSPEFKGKVCGLCGNFDDNAINDF
TTRSQSVVSDMLEFGNSWKLSPSCPDASVSKDPCTANPYRKSWAQKQCSIINSAAFSACHAHVEPAKYYE
ACVNDACACDSGGDCECFCTAVAAYAQACHEVGVCVSWRTPDICPLFCDYYNPEGQCEWHYQPCGAPCMR
TCQNPTGQCLQDLRGLEGCYPKCPPTAPIFDEGTMQCVSNCTVSPSPCRVNGKLYRPGTPIPSDENCYSC
VCTESGVNCTHDAGACVCTYNGQRYHPGDTIYHTTDGMGGCISAHCRDNGTIERIVDTCSSTSPPPPTTF
SFSTTLVMTSMQPSSTHSSPTPSVVYPGSPSKAVLTASSVSSVKTPETTSVLTTSTSASTLTMPACQEEC
LWSPWMDISRPGRGIDSGDFDTLENLHAHGYQICPVPKAVECRAEDNPGVPFHALQQHVECSTTVGLICY
NSDQVSGLCDNYQIKIQCCTPINCPTSTGPTQTTHLIVSRTSTMEDTTSSVPVTSTEHTYSTVASSPSTH
TPGPSPSSSVPSSSAPARSTPTPVSSTTVKTTLPTTSPMPEPTSATSSVSISTLGSTLASPEITHGCRKE
LCNWTDWIDGSYPEPGRSSGDFDTFVNLRAKGYKFCEKPWNVECRAQFFPNTPLQELGQDVTCSREVGLI
CLNKNQLPPICYNYEIRIECCTIVNICSTTSATTQPTSHGVSIKTKTNWITNTYSFSTENTSGHSTVINT
KTWVTGSTHTTPQPGTRPTPSIVSTQDTSTSSVQTDSTTSSSHTSSPNTGRVSTTHTTHTSSPPTGGTSP
TSTTHTSSPNTGGTSPTSTTHTSSPXTGGTS1
```